

UNIVERSIDAD DE COSTA RICA
SISTEMA DE ESTUDIOS DE POSGRADO

**ESTIMACIÓN DE UN MODELO DE REGRESIÓN DINÁMICA PARA LA
DEMANDA ELÉCTRICA RESIDENCIAL E INDUSTRIA DE COSTA RICA.**

Trabajo final de investigación aplicada sometido a la consideración de la Comisión del
Programa de Estudios de Posgrado en Estadística para optar por el grado y título de
Maestría Profesional en Estadística

MARCO VINICIO OTOYA CHAVARRIA

Ciudad Universitaria Rodrigo Facio, Costa Rica

2020

DEDICATORIA

A Dios por la salud y el tiempo para completar con éxito este proyecto académico y personal.

A mi compañera de maestría, de arduas jornadas estudio, de viajes y de vida; Melissa Cordero
Días Gracias por todo.

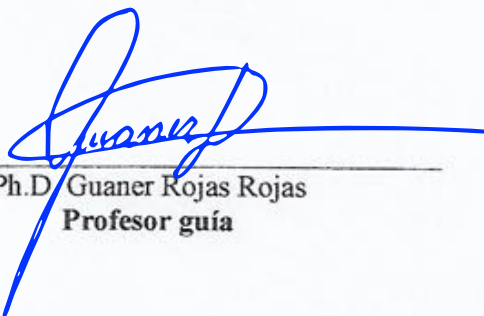
AGRADECIMIENTOS

Al Ph.D. Guaner Rojas Rojas por su colaboración, tiempo y dedicación en la dirección de mi trabajo final de graduación, así como su retroalimentación en este proceso y en los cursos en los que fui su alumno, los cuales fueron un espacio muy valioso de nuevos saberes y aprendizajes en mi proceso de formación en el área estadística.

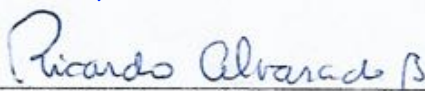
A mis lectores Ph.D Ricardo Alvarado Barrantes y Ph.D Leiner Vargas Alfaro por sus valiosos aportes y comentarios en este proceso.

Al Ph.D. Gilberth Brenes Camacho; M.Sc. Johnny Madrigal Pana y la M.Sc. Sandra Hernández, excelentes docentes y académicos.

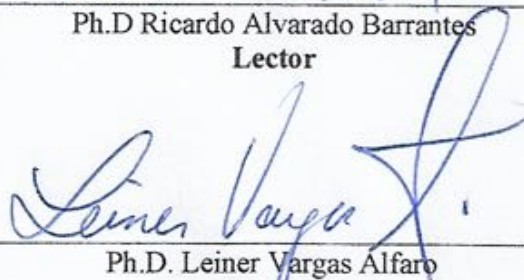
“Este trabajo final de investigación aplicada fue aceptado por la Comisión del Programa de Estudios de Posgrado en Estadística de la Universidad de Costa Rica como requisito parcial para optar al grado y título de Maestría Profesional en Estadística”




Ph.D. Guaner Rojas Rojas
Profesor guía



Ph.D. Ricardo Alvarado Barrantes
Lector



Ph.D. Leiner Vargas Alfaro
Lector



Marco Vinicio Otoy Chavarria
Sustentante

TABLA DE CONTENIDOS

DEDICATORIA	i
AGRADECIMIENTOS	i
TABLA DE CONTENIDOS	iii
RESUMEN	iv
SUMMARY	v
ÍNDICE DE TABLAS	vi
ÍNDICE DE CUADROS	vi
ÍNDICE DE GRÁFICOS	vi
ÍNDICE DE FIGURAS	vii
1. INTRODUCCIÓN	1
2.1 Problema	4
2.2 Objetivos	4
2.3 Justificación	4
2. MARCO TEÓRICO.....	7
3.1 Modelos Estadísticos.....	8
<i>a.1 Modelos de series de tiempo Box Jenkins.....</i>	<i>8</i>
<i>a.2 Modelo de retardos distribuidos</i>	<i>11</i>
<i>a.3 Modelos autorregresivos</i>	<i>14</i>
<i>a.4 Modelos vectoriales autoregresivos</i>	<i>15</i>
<i>a.5 Modelos de ecuaciones simultáneas</i>	<i>18</i>
3. METODOLOGÍA	22
4.1 Variables y Fuente de los datos.....	22
4.2 Modelo de series de tiempo	23
4.3 Modelo de Vectores Autorregresivos	25
4.4 Selección del modelo	28
4.5 Software a utilizar	28
4.6 Metodología de Estimación	30
4. RESULTADOS	32
5.1 Modelos de Series de tiempo.....	32
5.2 Modelos de Vectores Autorregresivos (VAR).....	53
5.3 Comparación de modelos	64
5. CONCLUSIONES	68
6. REFERENCIAS.....	72

RESUMEN

El presente trabajo realiza una estimación de la demanda eléctrica para el sector residencial e industrial (grandes consumidores) con fines de pronóstico, considerando la demanda total país y de los clientes del Instituto Costarricense de Electricidad. Tiene como objetivo comparar los resultados que pueden obtenerse mediante un modelo de series de tiempo utilizando en método de Box-Jenkins y un modelo de vectores autoregresivos que utiliza un sistema de ecuaciones con rezagos.

En cuanto a los modelos de series de tiempo, se realiza un proceso de identificación, estimación y verificación teniendo en cuenta los conocimientos en el tema, la teoría económica y econométrica, que además implica analizar la estacionariedad de la serie, evaluar el modelo, su capacidad predictiva entre otros. Adicionalmente, este proceso se compara con análisis automatizados que se realizan mediante la función `auto.arima` de R y la función `Tramo-Seat` de Eviews, lo que permite comparar para el caso de análisis la bondad de ajuste y predicción de los últimos dos métodos.

Disponer de modelos adecuados para la estimación de la demanda eléctrica resulta fundamental para fines regulatorios, por ejemplo, cuando se hace necesario definir una tarifa. Adicionalmente, permite determinar los niveles de inversión o necesidades del sistema en el corto y mediano plazo, siendo una herramienta fundamental para el correcto funcionamiento de los sistemas eléctricos.

Los resultados muestran como los modelos de series de tiempo ARIMA, al menos en este ejercicio, conllevan a mejores propósitos comparados con los modelos de vectores autoregresivos. Debe tenerse en cuenta que la elaboración de modelos de tipo ARIMA requieren un amplio análisis y entendimiento del comportamiento de la serie para identificar el modelo correcto. A su vez, los resultados obtenidos para este caso mediante la función `auto.arima` sugieren que esta función debe usarse con precaución y más como una guía o ayuda que como un instrumento para definir el modelo definitivo, dado que únicamente busca ajustar principalmente indicadores como el Criterio de Información Akaike.

SUMMARY

The research estimate of the electricity demand for the residential and industrial sector (large consumers) for forecasting purposes, considering the total demand for the country and for the clients of the Costa Rican Electricity Institute. The objective is to compare the results that can be obtained by time series model using the Box-Jenkins method and an autoregressive vector model that uses a system of equations with lags.

Regarding the time series models, a process of identification, estimation and verification is carried out taking into account the knowledge on the energy sector, the economic and econometric theory, which also involves analyzing the stationarity of the series, evaluating the model, its predictive capacity among others. Additionally, this process is compared with automated analyzes performed using the `auto.arima` function of R and the `Tramo-Seat` function of Eviews, which makes it possible to compare the goodness of fit and prediction of the last two methods for the case of analysis.

Having adequate models for estimating electricity demand is essential for regulatory purposes, for example, when it is necessary to define a tariff. Additionally, it makes it possible to determine the investment levels or needs of the system in the short and medium term, being a fundamental tool for the correct operation of electrical systems.

The results show how the ARIMA time series models, at least in this exercise, lead to better purposes compared to the autoregressive vector models. It should be noted that ARIMA-type modeling requires extensive analysis and understanding of series behavior to identify the correct model. In turn, the results obtained for this case using the `auto.arima` function suggest that this function should be used with caution and more as a guide or help than as an instrument to define the definitive model, since it only seeks to mainly adjust indicators such as the Akaike Information Criterion.

Índice de Tablas

Tabla 1. Descripción de variables y su fuente de información.....	23
---	----

Índice de Cuadros

Cuadro 1. Resultados de las pruebas de estacionariedad para la demanda eléctrica residencial.....	33
Cuadro 2. Resultados del modelo ajustado: ARIMA (0,1,1) (0,1,1) ₁₂	36
Cuadro 3. Resultados de las pruebas de estacionariedad para la demanda eléctrica residencial.....	40
Cuadro 4. Resultados del modelo ajustado: SARIMA (0,1,1) (1,1,0) ₁₂	42
Cuadro 5. Resultados de las pruebas de estacionariedad para la demanda eléctrica grandes consumidores.....	45
Cuadro 6. Resultados del modelo ajustado: SARIMA (1,1,1) (1,1,0) ₁₂	47
Cuadro 7. Resultados de las pruebas de estacionariedad para la demanda eléctrica grandes consumidores del ICE.....	50
Cuadro 8. Resultados del modelo ajustado: SARIMA (0,1,0) (0,1,1) ₁₂	52
Cuadro 9. Resultados del modelo VAR para la demanda eléctrica de grandes consumidores.....	54
Cuadro 10. Resultados del modelo VAR para la demanda grandes consumidores ICE.	57
Cuadro 11. Resultados del modelo VAR para la demanda eléctrica sector residencial.	60
Cuadro 12. Modelo VAR demanda eléctrica sector residencial ICE.	62
Cuadro 13. Error cuadrático medio para la proyección de la demanda eléctrica.	65

Índice de Gráficos

Gráfico 1. Costa Rica: Estacionalidad para el consumo eléctrico residencial. Período 2012-2019	32
Gráfico 2. ICE: Evolución de la demanda eléctrica residencial. Período 2012-2019.	39
Gráfico 3. ICE: Estacionalidad de la demanda eléctrica residencial. Período 2012-2019. ..	40
Gráfico 4. Costa Rica: Evolución de la demanda eléctrica para grandes consumidores. Período 2012-2019.	44
Gráfico 5. Costa Rica: Estacionalidad para el consumo eléctrico de grandes consumidores. Período 2012-2019	45
Gráfico 6. ICE: Evolución de la demanda eléctrica para grandes consumidores. Período 2012-2019.....	49
Gráfico 7. Costa Rica: Estacionalidad para el consumo eléctrico de grandes consumidores. Período 2012-2019.	50
Gráfico 8. Correlograma para los residuos del modelo.....	55
Gráfico 9. Correlograma para los residuos del modelo.....	58
Gráfico 10. Correlograma para los residuos del modelo.....	60

Gráfico 11. Correlograma para los residuos del modelo.	63
---	----

Índice de Figuras

Figura 1. Área de concesión y porcentaje de cobertura eléctrica por empresas distribuidora.	3
Figura 2. Etapas en la identificación y construcción de un modelo Box-Jenkins.	9
Figura 3. Resultados del Test Zivot and Andrews para determinar la presencia de raíz unitaria.	34
Figura 4. Funciones de autocorrelación simple y parcial para la serie original y diferenciada.	35
Figura 5. Residuos del modelo seleccionado (0,1,1)(0,1,1)[12], función de autocorrelación e histograma.	37
Figura 6. Resultados del Test Zivot and Andrews para determinar la presencia de raíz unitaria.	41
Figura 7. Funciones de autocorrelación simple y parcial para la serie original y diferenciada.	41
Figura 8. Residuos del modelo seleccionado (0,1,1)(1,1,0)[12], función de autocorrelación e histograma.	43
Figura 9. Resultados del Test Zivot and Andrews para determinar la presencia de raíz unitaria.	46
Figura 10. Funciones de autocorrelación simple y parcial para la serie original y diferenciada.	47
Figura 11. Residuos del modelo seleccionado (1,2,1)(1,1,0)[12], función de autocorrelación e histograma.	48
Figura 12. Resultados del Test Zivot and Andrews para determinar la presencia de raíz unitaria.	51
Figura 13. Funciones de autocorrelación simple y parcial para la serie original y diferenciada.	52
Figura 14. Residuos del modelo seleccionado (0,1,0)(0,1,1)[12], función de autocorrelación e histograma.	53
Figura 15. Prueba de estabilidad del modelo (OLS-CUSUM).	56
Figura 16. Prueba de estabilidad del modelo (OLS-CUSUM).	59
Figura 17. Prueba de estabilidad del modelo (OLS-CUSUM).	61
Figura 18. Prueba de estabilidad del modelo (OLS-CUSUM).	64
Figura 19. Proyección de la demanda eléctrica grades consumidores total país.	65
Figura 20. Proyección de la demanda eléctrica grades consumidores ICE.	66
Figura 21. Proyección de la demanda eléctrica sector residencial total país.	66
Figura 22. Proyección de la demanda eléctrica residencial del ICE.	67



UNIVERSIDAD DE
COSTA RICA

SEP Sistema de
Estudios de Posgrado

Autorización para digitalización y comunicación pública de Trabajos Finales de Graduación del Sistema de Estudios de Posgrado en el Repositorio Institucional de la Universidad de Costa Rica.

Yo, Marco Vinicio Otoya Chavarría, con cédula de identidad 1-1024-0363, en mi condición de autor del TFG titulado Estimación de un modelo de regresión dinámica para la demanda eléctrica residencial e industrial de Costa Rica.

Autorizo a la Universidad de Costa Rica para digitalizar y hacer divulgación pública de forma gratuita de dicho TFG a través del Repositorio Institucional u otro medio electrónico, para ser puesto a disposición del público según lo que establezca el Sistema de Estudios de Posgrado. SI ☒ NO * ☐

*En caso de la negativa favor indicar el tiempo de restricción: _____ año (s).

Este Trabajo Final de Graduación será publicado en formato PDF, o en el formato que en el momento se establezca, de tal forma que el acceso al mismo sea libre, con el fin de permitir la consulta e impresión, pero no su modificación.

Manifiesto que mi Trabajo Final de Graduación fue debidamente subido al sistema digital Kerwá y su contenido corresponde al documento original que sirvió para la obtención de mi título, y que su información no infringe ni violenta ningún derecho a terceros. El TFG además cuenta con el visto bueno de mi Director (a) de Tesis o Tutor (a) y cumplió con lo establecido en la revisión del Formato por parte del Sistema de Estudios de Posgrado.

INFORMACIÓN DEL ESTUDIANTE:

Nombre Completo: Marco Vinicio Otoya Chavarría

Número de Carné: B58437 Número de cédula: 1-1024-0363

Correo Electrónico: Marco.otoya.chavarria@una.ac.cr

Fecha: 21-10-2020 Número de teléfono: 8315-2771

Nombre del Director (a) de Tesis o Tutor (a): Ph.D Guaner Rojas Rojas

FIRMA ESTUDIANTE

Nota: El presente documento constituye una declaración jurada, cuyos alcances aseguran a la Universidad, que su contenido sea tomado como cierto. Su importancia radica en que permite abreviar procedimientos administrativos, y al mismo tiempo genera una responsabilidad legal para que quien declare contrario a la verdad de lo que manifiesta, puede como consecuencia, enfrentar un proceso penal por delito de perjurio, tipificado en el artículo 318 de nuestro Código Penal. Lo anterior implica que el estudiante se vea forzado a realizar su mayor esfuerzo para que no sólo incluya información veraz en la Licencia de Publicación, sino que también realice diligentemente la gestión de subir el documento correcto en la plataforma digital Kerwá.

1. INTRODUCCIÓN

La energía está presente en todos los aspectos de la vida moderna de un país desarrollado. Los distintos servicios que aporta la energía constituyen un factor muy relevante para la calidad de vida, no sólo actualmente, sino que ya llevan siéndolo desde hace muchas décadas, e incluso siglos (Club Español de la Energía, 2014). En este sentido, el sector eléctrico es fundamental para el crecimiento y el desarrollo de un país. Mediante la generación, transmisión, distribución y comercialización eléctrica se hace posible el desarrollo de una serie de actividades productivas y sociales, se estimula el comercio, la inversión y se contribuye a la mejora en la calidad de vida de la población.

La disponibilidad de energía eléctrica de manera continua y en condiciones de calidad, cobertura y confiabilidad, mejoran significativamente la disponibilidad y el acceso a la salud y la educación. Adicionalmente, al posibilitar el desarrollo de la industria y el comercio permite el desarrollo de la actividad económica y productiva del país, atrae la inversión nacional y extranjera generando posibilidades de empleo. La forma en que se planea suplir la demanda eléctrica también puede incidir positivamente en el desarrollo ambiental del país, cuando la mayor parte de la matriz eléctrica es mediante fuentes renovables.

En Costa Rica la generación eléctrica se realiza mediante la participación del sector público y el sector privado. El Instituto Costarricense de Electricidad (ICE) tiene la responsabilidad de asegurar el suministro eléctrico en el corto y largo plazo. Las cooperativas de electrificación rural y las empresas municipales también participan en la generación eléctrica para suplir parte de la demanda en su área de servicio. Por otra parte, la Ley N°. 7200 que autoriza la generación privada autónoma o paralela permite la generación privada con fuentes renovables, facultando al ICE a firmar contratos de compra de energía hasta por un máximo del 15% de la potencial del sistema eléctrico nacional.

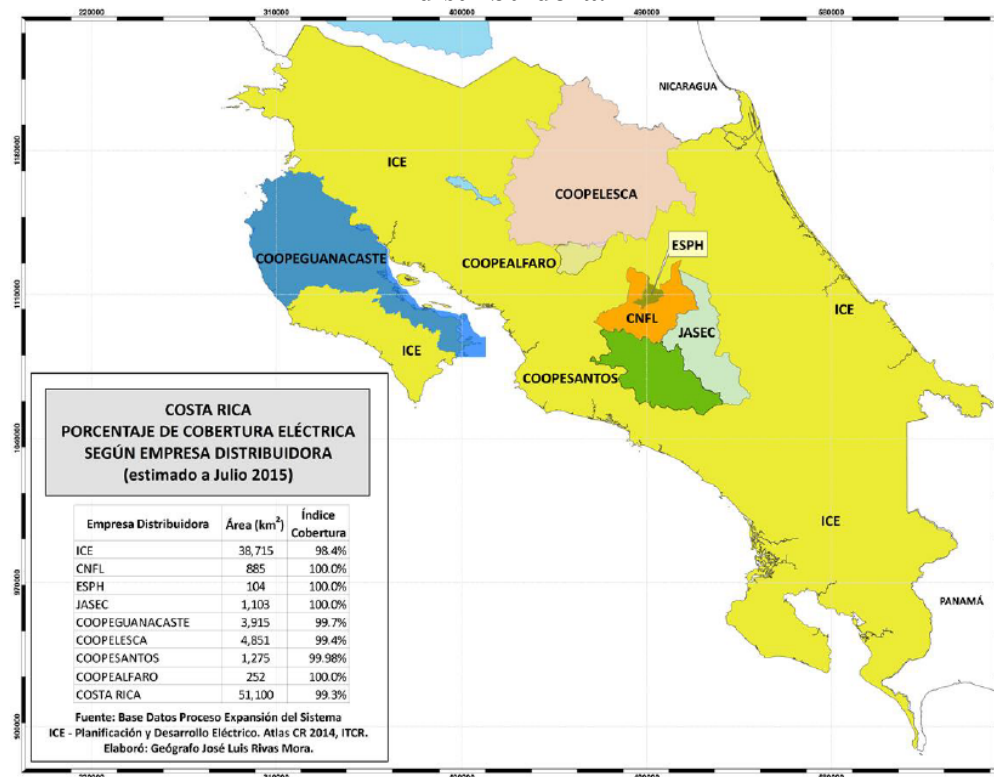
En el 2018 la generación eléctrica alcanzó los 12 367,52 GWh. El ICE generó el 61,25% de la electricidad, un 25,31% provino de los generadores privados y el restante 13,34% de las empresas municipales y las cooperativas de electrificación rural. Del total de la generación

un 74.7% fue generado con plantas hidroeléctricas, un 15,1% provino de fuentes eólicas, un 8,2% de la geotermia y el resto de las fuentes como la biomasa, energía solar y generación térmica. La producción eléctrica a partir de fuentes renovables conlleva beneficios sociales, económicos y ambientales, entre los que destacan bajas emisiones de gases de efecto invernadero, menor dependencia de combustibles fósiles, mayor aceptación social, generación de empleo y atracción de inversiones (Estado de la Nación, 2018.)

La transmisión de energía eléctrica es realizada exclusivamente por el ICE, siendo un monopolio natural en este segmento de la cadena. La distribución y comercialización de la electricidad se encuentra a cargo de diversas empresas, predominando una estructura de mercado oligopólica (pocas empresas), sin embargo, funcionan como un monopolio territorial dado que solo pueden distribuir y comercializar electricidad en la zona geográfica definida mediante concesión.

Como se observa en la Figura 1, el Instituto Costarricense de Electricidad brinda el servicio de distribución y comercialización eléctrica en la mayor parte del país, mientras que la Compañía Nacional de Fuerza y Luz (CNFL) es una subsidiaria del ICE y concentra sus servicios en el área metropolitana (GAM). La Empresa de Servicios Públicos de Heredia y la Junta Administradora del Servicio Eléctrico de Cartago, son empresas de capital municipal que tienen como área de influencia la GAM, particularmente algunos sectores de las provincias de Heredia y Cartago. Finalmente, existen seis cooperativas de electrificación rural que dan sus servicios en las zonas rurales: Cooperativa de Electrificación Rural de San Carlos (Coopelesca), Cooperativa de Electrificación Rural de Alfaro Ruíz (Coopealfaroruiz), Cooperativa de Electrificación Rural de Guanacaste (Coopeguancaste), y Cooperativa de Electrificación Rural de los Santos (COOPESANTOS R.L).

Figura 1. Área de concesión y porcentaje de cobertura eléctrica por empresas distribuidora.



Fuente: ICE. Plan de Expansión de la Generación Eléctrica 2016-2015.

Por el lado de la demanda, al 2018 se registraron un total de 1 753 895 abonados, que en función de la segmentación tarifaria pueden clasificarse como consumidores residenciales, industriales o general (comercio y servicios).

Las ventas de energía eléctrica por parte de las diferentes empresas distribuidoras a los sectores residencial e industrial reflejan la demanda de electricidad desde estos sectores hacia el sistema eléctrico nacional. Estudiar el comportamiento de la demanda para determinar sus componentes, estacionalidad, tendencia, y crecimiento futuro es fundamental en materia de planificación eléctrica, para el planeamiento de inversiones futuras, la confiabilidad y calidad del sistema y en la determinación de aspectos tarifarios. El análisis del comportamiento de la demanda permite, por ejemplo, identificar patrones estacionales y realizar pronósticos adecuados para la toma de decisiones.

2.1 Problema

La presente investigación buscar dar respuesta a las siguientes interrogantes: ¿Cómo se comporta la demanda eléctrica del sector residencial e industrial? ¿Existen patrones de comportamiento, tendencia o estacionalidad diferentes en ambos sectores? ¿Cuáles son los determinantes de la demanda eléctrica para el sector residencial e industrial? ¿Cuál es el modelo de regresión que proporciona el mejor ajuste para fines de pronóstico?

2.2 Objetivos

Objetivo General

Proponer un modelo de regresión dinámica para la predicción de la demanda eléctrica del sector residencial e industrial de Costa Rica con fines de pronóstico.

Objetivos Específicos

1. Describir el comportamiento en el tiempo de la demanda eléctrica en el sector residencial e industrial durante el período 2012-2018.
2. Identificar la relación dinámica existente entre el consumo eléctrico y variables económicas, sociales y ambientales potencialmente relevantes en el sector residencial e industrial.
3. Establecer un modelo de regresión dinámica con fines de pronóstico para los sectores residencial e industrial.

2.3 Justificación

Estudiar la demanda eléctrica resulta una actividad clave en materia regulatoria, para la planificación de inversiones y generar señales a los agentes económicos en materia de precios sobre el costo de producir energía eléctrica en diferentes momentos del día.

En materia regulatoria cada empresa distribuidora debe presentar a la Autoridad Reguladora de los Servicios Públicos, un estudio tarifario ordinario en el que se revisan y se determinan los costos de operación, el reconocimiento de inversiones y un rédito para el desarrollo o rentabilidad razonable. El conjunto de costos e inversiones a reconocer en un período vía tarifa deberá distribuirse entre sus clientes, por lo que es fundamental disponer de modelos de pronóstico que estimen de la mejor forma posible la demanda potencial. La subestimación o sobre estimación de la demanda genera problemas tanto para el operador como para los clientes finales, provocando un desequilibrio financiero o un exceso de recursos para las empresas distribuidoras o bien una tarifa más baja o mucho más elevada de lo que debería ser en el caso de los usuarios finales.

Los pronósticos de demanda son la llave para un buen planeamiento de la operación de los sistemas eléctricos, ya que dependiendo del horizonte de tiempo a evaluar, pueden ayudar a determinar si en algún momento se llegará a producir un déficit de la capacidad de generación y en consecuencia, puede ser conveniente considerar la construcción de nuevas centrales de generación de energía, expansión del sistema de transmisión o impulsar políticas de ahorro energético, o de otra manera, puede indicar que hay un exceso de capacidad de generación (Orellana J, 2012).

La estimación correcta de las variaciones que puedan suceder en el corto, mediano y largo plazo en términos de la demanda, también contribuye al planeamiento de las inversiones necesarias para garantizar la respectiva oferta. Por ejemplo, un faltante de capacidad de generación en el país tendría un impacto muy grave, habida cuenta de los altos costos que implicaría para la sociedad costarricense un desabastecimiento; por su parte, inversiones excesivas aumentan el costo de la energía y causan un daño a la economía nacional en su conjunto (ICE, 2017).

Con respecto a las inversiones, las estimaciones de demanda también pueden contribuir a direccionarlas de acuerdo con el crecimiento esperado de esta. Particularmente, el instrumento utilizado para la planificación de la oferta eléctrica en el país se realiza mediante

el Plan de Expansión de la Generación, que tiene como propósito proponer una estrategia de desarrollo del sistema de generación del país a largo plazo y que está a cargo del ICE.

El conocimiento de la demanda y su propósito también puede permitir elaborar estrategias e instrumentos económicos para su gestión. Esto puede hacer posible establecer estructuras tarifarias que incentiven o desincentiven el consumo eléctrico en ciertos momentos del día, por ejemplo, en función del costo de la generación, enviando señales a los agentes económicos para la toma de decisiones.

La estimación de un modelo de demanda eléctrica se vuelve sumamente relevante en un contexto en donde las tecnologías disruptivas como la generación distribuida y las redes inteligentes podrían modificar los patrones de consumo de los distintos actores. Así mismo, la política pública ha influido en los últimos años en el precio de algunos bienes sustitutos como lo es el gas licuado de petróleo, lo que está provocando un desplazamiento de la demanda por electricidad hacia el gas.

Los factores ambientales también son un aspecto importante para considerar. En el país, la mayor parte de la energía generada proviene de fuentes renovables en particular hidroeléctrica, las cuales durante los primeros meses del año se ven afectadas por la época seca, lo que se traduce en una menor disponibilidad del recurso para generación y mayores tarifas; en especial cuando se requiere generar energía con combustibles fósiles.

Tradicionalmente, se han utilizado modelos de series de tiempo del tipo Box Jenkins (ARIMA) para estimaciones de demanda eléctrica; aunque también se utilizan técnicas de promedios móviles, suavización exponencial (simple, Croston, Holt, Winters y variaciones), y más recientemente técnicas como redes neuronales. Estos modelos por lo general toman en cuenta el comportamiento en el tiempo de la variable a modelar para predecir su comportamiento futuro. Sin embargo, en un contexto dinámico es necesario disponer de modelos dinámicos que estudien la relación entre la variable endógena y las variables explicativas, pero en diferentes momentos de tiempo. Entre este tipo de modelos se

encuentran los modelos de rezagos distribuidos y modelos de ecuaciones funcionales, entre otros.

2. MARCO TEÓRICO

La función de demanda de un bien energético expresa las relaciones que existen entre las cantidades consumidas de un bien por unidad de tiempo y los distintos valores que asumen los determinantes de esa demanda (Leiva C; 2009). Diversos son los factores que tienden a influir en los mercados de energía y, por tanto, en la demanda eléctrica entre los que se pueden indicar la tarifa, el precio de otros energéticos sustitutos como el Gas Licuado del Petróleo, y el ingreso aproximado mediante algún indicador como por ejemplo el producto interno bruto. Adicionalmente, otros factores pueden ser considerados como, por ejemplo, la elasticidad precio de la demanda en el corto plazo, los riesgos regulatorios, la asimetría de información a la que son sometidos los agentes económicos, la estructura y naturaleza de la oferta de mercado (Barrientos J y Toro Mónica, 2017), temperatura y día de la semana (Carolina García y Carlos Mate, 2010).

De acuerdo con (Carolina García y Carlos Mate, 2010) la mayoría de los métodos de pronóstico en el sector eléctrico utilizan técnicas estadísticas o herramientas de inteligencia artificial. Los estudios iniciales basados en modelos estadísticos utilizan por ejemplo modelos de medias móviles (ARIMA) (Abdel-Aal and Al-Garni, 1997; Saab et al., 2001) o modelos basados en regresión (Mohgram and Rahman, 1989; Papalexopoulos and Hesterberg, 1990).

Otros modelos basados técnicas de cointegración han utilizado un enfoque de datos de panel para estudiar la demanda eléctrica (Alaa El-Shazly; 2013), en este caso el modelo estimado permite la heterogeneidad transversal en un marco dinámico incorporando información sobre ingresos y precios relevantes de bienes nacionales y extranjeros. Por otra parte, Vilar J, et al; (2018) mediante el uso de datos funcionales estimó dos métodos para obtener intervalos de predicción para la demanda de electricidad y precio con pronósticos puntuales para el día siguiente. El primer método empleó un modelo autorregresivo no paramétrico y el segundo

utiliza un modelo semiparamétrico lineal parcial, en el que las covariables escalares exógenas se incorporan de manera lineal.

Con el fin de pronosticar la demanda mensual de energía eléctrica por hora en España durante 2 años (Carolina García y Carlos Mate, 2010) realizaron una comparación entre un nuevo enfoque de previsión considerando los modelos de predicción de vectores autorregresivos (VAR) aplicados a series temporales de intervalo (ITS) y el iMLP, Modelo de capa perceptrón adaptado a datos de intervalo.

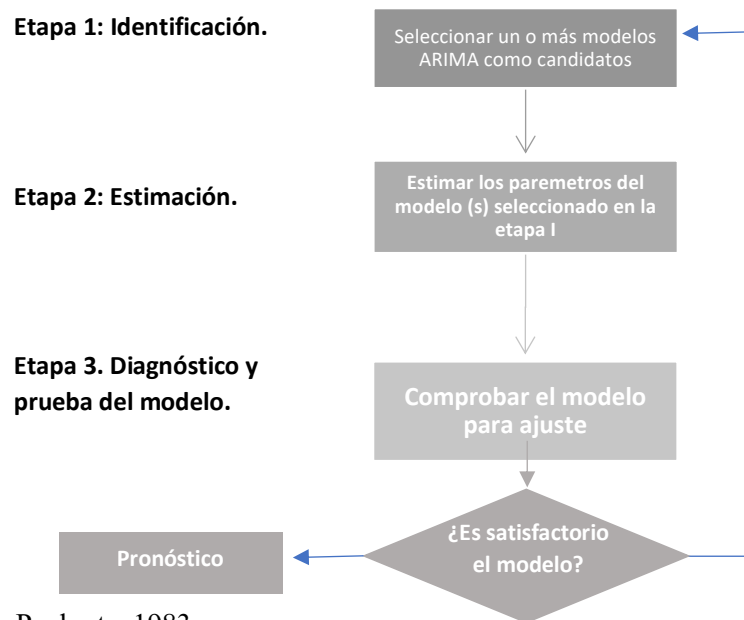
3.1 Modelos Estadísticos

a.1 Modelos de series de tiempo Box Jenkins

Las observaciones en una serie de tiempo pueden estar estadísticamente relacionadas con otras observaciones en la misma serie. El objetivo es encontrar una forma de establecer esa relación estadística, es decir un modelo que describa como se relacionan entre sí las observaciones en una única serie de tiempo (Pankratz,1983).

Un modelo ARIMA es muestra como una variable temporal (Z_t) esta relacionada con sus valores pasados ($Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots$). Si se considera que el siguiente modelo $Z_t = C + \Phi_1 Z_{t-1} + a_1$, es un modelo ARIMA que implica que Z_t esta relacionado con su propio valor pasado Z_{t-1} , C es una constante y el termino Φ_1 es el coeficiente fijo cuyo valor determina la relación entre Z_t y Z_{t-1} . El término a_1 es el término de error probabilístico. Todos los componentes del lado derecho de la ecuación (exceptuando a_1) son componentes de Z_t , por tanto, representan la parte predictiva de Z_t mientras que a_1 es el término de error estocástico, es decir un residuo que no puede ser predicho mediante el modelo ARIMA. Box and Jenkins propone un procedimiento en tres etapas para ajustar un modelo ARIMA:

Figura 2. Etapas en la identificación y construcción de un modelo Box-Jenkins.



Fuente: Pankratz, 1983.

Etapa 1: Identificación. La etapa de identificación utiliza el instrumental gráfico para medir la correlación entre las observaciones en una serie de datos, para ello se emplea una estimación de la función de autocorrelación (AFC) y una estimación de la función de autocorrelación parcial (PACF). Estas funciones miden las relaciones estadísticas dentro de una serie de datos (estadísticamente ineficiente). Sin embargo, son útiles para proporcionar una idea de los patrones que siguen las observaciones disponibles.

Un segundo paso en la etapa de identificación es resumir las relaciones estadísticas dentro de una serie de datos de forma más compacta de lo que se hace mediante el ACF y PACF. Box Jenkins sugieren una serie de modelos teóricos algebraicos entre los que se pueden seleccionar. Cualquiera que sea el modelo para seleccionar en la etapa de identificación, se elige solo de manera tentativa, es decir, es solo un candidato para el modelo final. Para elegir un modelo final, se continúa con las siguientes etapas, incluso puede ser necesario regresar a la etapa de identificación si el modelo considerado provisionalmente resulta inadecuado.

Etapa 2: Se obtiene una estimación precisa de los coeficientes del modelo seleccionado en la etapa de identificación. Esta etapa proporciona señales con respecto a la idoneidad del modelo, en particular si los coeficientes estimados no satisfacen ciertas condiciones matemáticas, ese modelo es rechazado.

Etapa 3. Box y Jenkins sugieren algunas pruebas de diagnóstico que ayudan a determinar si un modelo es estadísticamente adecuado, si el modelo falla estas pruebas es rechazado. Si el modelo es rechazado se debe regresar a la etapa de identificación y repetir el proceso hasta encontrar un modelo definitivo.

La etapa de estimación y validación del diagnóstico provee señales de alerta indicando cuando y como el modelo puede ser reformulado. Con estas señales el modelo continúa con la etapa de re-identificación, re-estimación y re-validación hasta obtener un modelo satisfactorio de acuerdo con varios criterios.

Diferenciación de las series

La mayoría de las series económicas son no estacionarias, esto implica que la media y la varianza cambian en el tiempo, es decir no son constantes. En estos casos es necesario aplicar alguna técnica para transformar las series en procesos estacionarios, comúnmente este proceso es llamado diferenciación de las series. La diferenciación es una operación relativamente simple que involucra Calcular cambios sucesivos en los valores de una serie de datos.

Para la diferenciación de las series de tiempo se calculan los cambios período a periodo $w_t = z_t - z_{t-1}$, en el caso de una diferenciación. Para diferenciar una serie dos veces se calculan los cambios sobre las primeras diferencias $w_t = (z_t - z_{t-1}) - (z_{t-1} - z_{t-2})$. En la práctica la primera diferencia se requiere con bastante frecuencia, mientras que la segunda diferenciación solo es necesaria de vez en cuando. Una tercera diferenciación casi nunca es necesaria.

a.2 Modelo de retardos distribuidos

La dependencia de una variable Y (variable dependiente) respecto a otra y otras variables X (variables explicativas) raramente es instantánea. Muy frecuentemente, Y responde a X en un lapso de tiempo, el cual se denomina rezago (Gujarati, 2010).

En los modelos de retardos distribuidos de orden S se explicita la dinamicidad mediante a la presencia de variables exógenas retardadas dentro de los regresores, es decir, el efecto de una causa dada se propaga o se distribuye en el tiempo. El orden del modelo de retardos distribuidos es igual al número de retardos asociados a las variables exógenas. El modelo puede ser descrito mediante la siguiente ecuación:

$$Y_t = \alpha + \Phi_0 X_t + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \cdots + \Phi_k X_{t-k} + u_t$$

Siendo un modelo de rezagos distribuidos con un rezago finito de k periodos. El coeficiente Φ_0 se conoce como multiplicador de corto plazo o de impacto, porque el cambio en el valor medio de Y sigue un cambio unitario en X en el mismo período. Si el cambio en X se mantiene al mismo nivel desde el principio, entonces $(\Phi_0 + \Phi_1)$ da el cambio en el valor medio de Y en el periodo siguiente, lo que se puede extender sucesivamente en cada período, dichas sumas parciales se denominan multiplicadores intermedios. El multiplicador de rezagos distribuidos de largo plazo, después de k periodos sería:

$$\sum_{i=0}^k \Phi_i = \Phi_0 + \Phi_1 + \Phi_2 + \cdots + \Phi_k = \Phi$$

Se define $\Phi_i = \frac{\Phi_i}{\sum \Phi_i} = \frac{\Phi_i}{\Phi}$, se obtiene el Φ_i “estandarizado” que dan la proporción del impacto en el largo plazo.

Diversas razones explican la naturaleza de los rezagos, entre ellas se pueden considerar las razones psicológicas, que se dan como resultado de la fuerza del hábito (inercia), las personas o empresas no cambia sus hábitos de consumo o producción de inmediato tras un cambio en

el entorno económico; asimismo, se puede no saber si un cambio es “permanente” o “transitorio”, de lo que dependerá la reacción final, por tanto la información incompleta también puede explicar en medida la existencia de rezagos. Adicionalmente, pueden mencionarse razones de tipo tecnológico o institucional. En el primer caso, a pesar de un cambio tecnológico o una reducción en el precio de algún insumo, dada la tecnología, puede no hacer técnicamente factible un ajuste de forma inmediata. En el segundo caso, por ejemplo, las razones contractuales pueden limitar los ajustes ante variaciones de las variables económicas en el corto plazo, por lo que estos ajustes se distribuirán en el tiempo.

Los modelos de rezagos distribuidos pueden ser de tipo finito o infinito. En el modelo de rezagos distribuidos finitos se permite que una o más variables influyan en “y” en forma rezagada, especificando la longitud del rezago k. Por su parte, el modelo de rezagos distribuidos infinitos, no se define la longitud del rezago, es decir, que tan atrás se desea ir.

Estimación ad hoc de un modelo de rezagos distribuidos:

Como se supone que la variable explicativa X_t es no estocástica (o por lo menos no correlacionada con el término de perturbación u_t), igualmente son no estocásticas X_{t-1} , X_{t-2} , y así sucesivamente. Por consiguiente, en principio, es aplicable el método de mínimos cuadrados ordinarios. Alt (1942) y Tinbergen (1949) sugieren que para estimar un modelo de rezagos distribuidos se procede secuencialmente, es decir, primero la regresión Y_t sobre X_t , luego la de Y_t sobre X_t y X_{t-1} , después la regresión de Y_t sobre X_t , X_{t-1} , y X_{t-2} y así sucesivamente. Este procedimiento secuencial se detiene cuando los coeficientes de regresión de las variables rezagadas empiezan a ser estadísticamente insignificantes o el coeficiente de por lo menos una variable cambia su signo de positivo a negativo, o viceversa.

La estimación ad hoc presenta algunas desventajas que deben ser consideradas. No hay una guía a prior sobre la longitud máxima que debe tener el rezago, la estimación de rezagos sucesivos limita la disponibilidad de grados de libertad y los valores sucesivos suelen estar altamente correlacionados generando problemas de multicolinealidad.

Método de Koyck, expectativas adaptativas y de ajuste parcial para los modelos de rezagos distribuidos

Se han diseñado algunos métodos abreviados para la estimación de modelos con rezagos distribuidos: los mecanismos de Koyck, de expectativas adaptativas y de ajuste parcial. El primero es un método puramente algebraico y los otros dos se basan en principios económicos.

El método de Koyck supone iniciar con un modelo de rezagos distribuidos finitos, si todas las β tienen el mismo signo, supone que se reducen geométricamente:

$$\beta_k = \beta_0 \lambda^k \quad k = 0, 1, \dots$$

Donde λ , tal que $0 < \lambda < 1$, se conoce como la tasa de descenso, o de caída, del rezago y donde $1 - \lambda$ se conoce como la velocidad de ajuste. Cada coeficiente β sucesivo es numéricamente inferior a cada β anterior, por lo que a medida que se retorna al pasado distante, el efecto de ese rezago sobre Y_t se reduce progresivamente.

Dado que el modelo anterior es meramente algebraico se han desarrollado dos enfoques previstos de soporte teórico. El modelo de expectativas adaptativas supone que los “agentes económicos adaptarán sus expectativas a la luz de la experiencia pasada y que, en particular, aprenderán de sus errores”, se establece que las expectativas se corrigen cada periodo por una fracción γ de la brecha entre el valor actual de la variable y su valor esperado anterior. Por otra parte, el modelo de ajuste parcial se desarrolló en el marco del análisis de las existencias y nivel deseado de capital; Marc Nerlove inició este modelo suponiendo un nivel de existencias de capital óptimo o a largo plazo sugerido para proporcionar una producción dada en unas ciertas condiciones de tecnología, tasa de interés.

El modelo de ajuste parcial se basa en que el cambio real de capital ($Y_t - Y_{t-1}$), en un periodo de tiempo t , es una fracción ρ del cambio deseado ($Y_t^* - Y_{t-1}$) durante ese mismo periodo. Por tanto $Y_t - Y_{t-1} = \rho(Y_t^* - Y_{t-1})$, donde ρ , tal que $0 < \rho \leq 1$, se conoce como coeficiente

de ajuste y donde $Y_t - Y_{t-1}$ es igual al cambio real, y $\rho(Y_t^* - Y_{t-1} - Y_{t-1})$ es igual al cambio deseado.

Una característica única de los modelos de Koyck, de expectativas adaptativas y de ajuste parcial es que todos son autorregresivos por naturaleza, es decir, el valor o valores de la variable regresada aparecen como una de las variables explicativas. Para estimar los modelos de Koyck y de expectativas adaptativas consistentemente, el método más común es el método de variables instrumentales. La variable instrumental es una variable representante para la variable regresada rezagada, pero con la propiedad de que no está correlacionada con el término de error.

a.3 Modelos autorregresivos

Tanto los modelos de Koyck, como de expectativas tiene una forma común y de ajuste parcial tienen la siguiente forma común:

$$Y_t = \alpha(1 - \lambda) + \beta_0 X_t + \lambda Y_{t-1} + v_t \quad \rightarrow \quad \text{Koyck}$$

$$Y_t = \gamma\beta_0 + \gamma\beta_1 X_t + (1 - \gamma)Y_{t-1} + [\mu_t - (1 - \gamma)\mu_{t-1}]$$

$$\rightarrow \quad \text{Expectativas adaptativas}$$

$$Y_t = \delta\beta_0 + \delta\beta_1 X_t + (1 - \delta)Y_{t-1} + \delta\mu_t \quad \rightarrow \quad \text{Ajuste parcial}$$

Todos estos modelos tienen la siguiente forma común:

$$Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 Y_{t-1} + v_t$$

Los modelos de Koyck son autorregresivos por naturaleza, es decir, se basa en la estimación mediante los rezagos de la variable dependiente, además de permitir el uso de un número restringido de rezagos en lugar de uno finito; por lo que se requiere un método de estimación diferente de Mínimos Cuadrados Ordinarios (MCO) dada la presencia de variables explicativas estocásticas y la posibilidad de correlación serial. El hecho de utilizar un modelo diferente a MCO se relaciona con la necesidad de demostrar que la variable explicativa

estocástica Y_{t-1} está distribuida independientemente del término de perturbación v_t , para lo cual deben conocerse las propiedades de este término de perturbación. Suponiendo que el término de perturbación original u_t satisface todos los supuestos clásicos, como $E(u_t) = 0$ y $var(u_t) = \sigma^2$ y no autocorrelación, sin embargo, v_t puede no heredar todas estas propiedades, estando serialmente correlacionada en el modelo de Koyck.

El problema surge dado que, si una variable explicativa en un modelo de regresión está correlacionada con el término de perturbación estocástico, los estimadores MCO están sesgados y son inconsistentes; es decir, aunque el tamaño de la muestra aumente indefinidamente, los estimadores no se aproximarán a sus valores poblacionales verdadero

Para estimar los modelos de Koyck y de expectativas adaptativas consistentemente, el método más común es el método de variables instrumentales. La variable instrumental es una variable representante para la variable regresada rezagada pero con la propiedad de que no está correlacionada con el término de error.

a.4 Modelos vectoriales autoregresivos

El modelo de vectores autoregresivos (VAR) permite caracterizar las interacciones simultaneas entre un conjunto de variables. Un VAR es un modelo de ecuaciones simultaneas formado por un sistema de ecuaciones de forma reducida sin restringir (Novales A; 2017). Un conjunto de ecuaciones de forma reducida implica que los valores contemporáneos de las variables del modelo no aparecen como variables explicativas en ninguna de las ecuaciones, en este sentido el conjunto de variables explicativas de cada ecuación esta constituido por un bloque de retardos de cada una de las variables del modelo. Que sean ecuaciones no restringidas significa que aparece en cada una de ellas el mismo grupo de variables explicativas.

También pueden incluirse como variables explicativas algunas variables de naturaleza determinista, como una posible tendencia temporal, variables ficticias estacionales, o una variable ficticia de tipo impulso, que sirve para llevar a cabo un análisis de intervención en el sistema, o bien, incluirse como explicativa una variable, incluso en valor contemporáneo,

que pueda considerarse exógena respecto a las variables que integran el modelo VAR (Novales A; 2017).

De acuerdo con Greene W. (1999) los VAR simplemente son una forma reducida sobreajustada de algún modelo de ecuaciones simultaneas. El sobre ajuste resulta de la posible inclusión de más retardos que los que serían apropiados en el modelo original. Una de las virtudes de los VAR es que obvia la decisión sobre que variables contemporáneas son exógenas. Tiene solamente variables retardadas en la parte derecha. La proliferación de parámetros en los VAR es frecuentemente citada como la principal desventaja de su uso.

El modelo VAR es muy útil cuando existe evidencia de simultaneidad entre un grupo de variables, y que sus relaciones se transmiten a lo largo de un determinado número de períodos. Al no imponer ninguna restricción sobre la versión estructural del modelo, no se incurre en los errores de especiación que dichas restricciones pudieran causar al ejercicio empírico. De hecho, la principal motivación detrás de los modelos VAR es la dificultad en identificar variables como exógenas, como es preciso hacer para identificar un modelo de ecuaciones simultáneas (Novales A; 2017).

Un VAR es un modelo de ecuaciones simultáneas formado por un sistema de ecuaciones de forma reducida sin restringir. Que sean ecuaciones de forma reducida quiere decir que los valores contemporáneos de las variables del modelo no aparecen como variables explicativas en ninguna de las ecuaciones. Por el contrario, el conjunto de variables explicativas de cada ecuación está constituido por un bloque de retardos de cada una de las variables del modelo. Que sean ecuaciones no restringidas significa que aparece en cada una de ellas el mismo grupo de variables explicativas.

Un modelo VAR en su forma simple puede escribirse de la siguiente manera:

$$Y_{1t} = \alpha_{10} + \alpha_{11}Y_{1t} + \alpha_{12}Y_{1t-1} + \alpha_{13}Y_{2t-1} + \gamma_1Z_t + \varepsilon_{1t}$$

$$Y_{2t} = \alpha_{20} + \alpha_{21}Y_{1t} + \alpha_{22}Y_{1t-1} + \alpha_{23}Y_{2t-1} + \gamma_2Z_t + \varepsilon_{2t}$$

Donde Y_{1t} , Y_{2t} son variables estacionarias, y ε_{1t} , ε_{2t} son procesos ruido blanco con esperanza cero, varianzas $\sigma_{\varepsilon_1}^2$, $\sigma_{\varepsilon_2}^2$ y covarianza σ_{12} . Este es un ejemplo de ecuaciones simultáneas con dos variables endógenas, Y_{1t} , Y_{2t} , y un vector z_t de variables exógenas. Un shock inesperado sobre Y_{2t} , en la forma de un valor no nulo de la innovación estructural ε_{2t} , afecta directamente a Y_{2t} , pero también incluye sobre Y_{1t} a través de la presencia de Y_{2t} como variable explicativa en la primera ecuación. Además, ese efecto se propaga en el tiempo, debido a la presencia de los valores retardados de ambas variables como variables explicativas. La identificación del número de retardos adecuados forma parte del análisis de especificación del modelo.

El modelo estructural puede incorporar asimismo un vector de variables explicativas exógenas z_t en cada ecuación, que puede aparecer asimismo con rezagos, como lo puede ser una tendencia determinística o variables ficticias estacionales. La representación matricial del modelo estructural de primer orden se presenta de la siguiente manera:

$$BY_t = \Gamma_0 + \Gamma_1 Y_{t-1} + Gz_t + \varepsilon_t$$

En el caso del modelo de dos ecuaciones, las matrices de esta representación son:

$$B = \begin{pmatrix} 1 & -\alpha_{21} \\ -\alpha_{21} & 1 \end{pmatrix}; \Gamma_0 = \begin{pmatrix} \alpha_{10} \\ \alpha_{20} \end{pmatrix}; \Gamma_1 = \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{22} & \alpha_{23} \end{pmatrix}; G = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$$

El modelo estructural VAR presenta dos dificultades para su estimación: a) la simultaneidad, al aparecer cada una de las dos variables como variable explicativa en la ecuación de la otra, lo que genera inconsistencia del estimador MCO, b) si los términos de error tuviesen autocorrelación, las estimaciones.

MCO serían inconsistentes, al tratarse de un modelo dinámico. En cualquiera de estos casos se incumpliría la condición de ausencia de correlación entre cada una de las variables explicativas y el término de error de una misma ecuación $E(x_{1t}\varepsilon_{1t}) = 0$, $E(x_{2t}\varepsilon_{2t}) = 0$, lo que garantizaría la consistencia del estimador de mínimos cuadrados. Pero en este modelo:

$x_{1t} = (Y_{2t}, Y_{1t-1}, Y_{2t-1}, z_t)$, y $x_{2t} = (Y_{1t}, Y_{1t-1}, Y_{2t-1}, z_t)$, por lo que las condiciones necesarias para la consistencia se incumplen. Por tanto, habría que esperar que dicho estimador fuese en este caso inconsistente.

La primera dificultad podría resolverse estimando por variables instrumentales, siempre que contemos con instrumentos adecuados, lo cual no es sencillo de justificar. Por eso, para evitarla, transformamos el modelo. Pero la segunda dificultad persistirá, y se debe resolver tratando de ampliar la estructura dinámica del modelo hasta lograr que los términos de error carezcan de autocorrelación.

El modelo vectorial autoregresivo (VAR) tiene un orden, que es el número de retardos con que las variables entran en cada ecuación.

a.5 Modelos de ecuaciones simultáneas

Existen situaciones en las que la relación causa-efecto en un sentido, o unidireccional, no es relevante. Esto sucede cuando Y está determinada por las X y algunas X , a su vez, están determinadas por Y . Por tanto, hay una relación en dos sentidos, o simultánea, entre Y y (algunas) X , que hace dudar del valor de la distinción entre variables dependientes y explicativas. Un modelo de ecuaciones simultáneas lo que hace es reunir un conjunto de variables que se determinen simultáneamente mediante el conjunto restante de variables. En dichos modelos hay más de una ecuación, una para cada una de las variables mutuamente, o conjuntamente, dependientes o endógenas.

A diferencia de los modelos uniecuacionales, en los modelos de ecuaciones simultáneas no es posible estimar los parámetros de una ecuación aisladamente sin tener en cuenta la información proporcionada por las demás ecuaciones en el sistema. Para que un sistema sea completo, se requiere que el número de ecuaciones iguale al número de variables endógenas (Greene W, 1999).

Problema de identificación

Una ecuación particular se encuentra bien especificada cuando las estimaciones numéricas de los parámetros de una ecuación estructural pueden obtenerse de los coeficientes estimados en forma reducida, en caso contrario, la ecuación esta no identificada o subidentificada. Por su parte, una ecuación identificada puede estar exactamente identificada si pueden obtenerse valores numéricos únicos de los parámetros estructurales, o bien, puede estar sobreidentificada si puede obtenerse más de un valor numérico para algunos de los parámetros de las ecuaciones estructurales.

El problema de identificación aparece porque diferentes conjuntos de coeficientes estructurales pueden ser compatibles con el mismo conjunto de información. Lo anterior implica que una ecuación en una forma reducida dada puede ser compatible con diferentes ecuaciones estructurales o con diferentes modelos.

Las ecuaciones en forma reducida permiten determinar la identificación de una ecuación en un sistema de ecuaciones simultáneas. Sin embargo, las denominadas condiciones de orden y de rango de identificación permiten realizar este proceso de manera más eficiente; sea:

M = número de variables endógenas en el modelo.

m = número de variables endógenas en una ecuación dada.

K = número de variables predeterminadas en el modelo, incluyendo el intercepto.

k = número de variables predeterminadas en una ecuación dada.

Condición de orden para la identificación

Una condición necesaria (pero no suficiente) para la identificación, conocida como la condición de orden. El término orden refiere al orden de una matriz, es decir, el número de filas y de columnas que contiene. Esta condición puede expresarse en dos formas diferentes pero equivalentes, de la siguiente manera (las condiciones necesaria y suficiente para la identificación se presentan más adelante):

- En un modelo de M ecuaciones simultáneas, para que una ecuación esté identificada debe excluir *al menos* $M - 1$ variables (endógenas y predeterminadas) que aparecen

en el modelo. Si excluye exactamente $M - 1$ variables, la ecuación está exactamente identificada. Si excluye más de $M - 1$ variables, estará sobreidentificada.

- En un modelo de M ecuaciones simultáneas, para que una ecuación esté identificada, el número de variables predeterminadas excluidas de esa ecuación no debe ser menor que el número de variables endógenas incluidas en la ecuación menos 1, es decir,

$$K - k \geq m - 1$$

Si $K - k = m - 1$, la ecuación está exactamente identificada, pero si $K - k > m - 1$, estará sobreidentificada.

Como lo indican los ejemplos anteriores, la identificación de una ecuación en un modelo de ecuaciones simultáneas es posible si dicha ecuación excluye una o más variables que están presentes en otras partes del modelo. Esta situación se conoce como *criterio de exclusión* (de variables), o *criterio de cero restricciones* (se supone que los coeficientes de las variables que no aparecen en una ecuación tienen valor de cero). Este criterio es el más utilizado para asegurar o determinar la identificación de una ecuación. Observe que el criterio de cero restricciones está basado en expectativas *a priori* o teóricas acerca de la ausencia de ciertas variables en una ecuación dada. Depende del investigador señalar claramente la razón por la cual espera que ciertas variables aparezcan en algunas ecuaciones y en otras no.

Condición de rango para la identificación

La condición de orden es una condición necesaria pero no suficiente para la identificación; es decir, aun si se cumple, puede suceder que una ecuación no esté identificada. Aun si una ecuación cumple la condición de orden $K - k \geq m - 1$, puede no estar identificada porque las variables predeterminadas excluidas de esa ecuación, pero presentes en el modelo, quizá no todas sean independientes de manera que tal vez no exista una correspondencia uno a uno entre los coeficientes estructurales (las β) y los coeficientes en forma reducida. O sea, probablemente no sea posible estimar los parámetros estructurales a partir de los coeficientes en la forma reducida. Por consiguiente, se requiere una condición que sea tanto necesaria

como suficiente para la identificación. Ésta es la condición de rango para la identificación, la cual puede expresarse en los siguientes términos:

- En un modelo que contiene M ecuaciones en M variables endógenas, una ecuación está identificada si y sólo si puede construirse *por lo menos* un determinante diferente de cero, de orden $(M - 1)(M - 1)$, a partir de los coeficientes de las variables (endógenas y predeterminadas) excluidas de esa ecuación particular, pero incluidas en las otras ecuaciones del modelo.

Como lo muestra el análisis anterior, la condición de rango dice si la ecuación bajo consideración está identificada o no, en tanto que la condición de orden expresa si dicha ecuación está exactamente identificada o sobreidentificada.

Prueba de simultaneidad

Si no hay ecuaciones simultáneas, o presencia del problema de la simultaneidad, MCO producen estimadores consistentes y eficientes. Por otra parte, si hay simultaneidad, los estimadores de MCO no son ni siquiera consistentes. En presencia de simultaneidad, los métodos de mínimos cuadrados en dos etapas (MC2E) y de variables instrumentales (VI) producirán estimadores consistentes y eficientes. Extrañamente, si se aplican métodos alternos cuando de hecho no hay simultaneidad, éstos producen estimadores que son consistentes, pero no eficientes (es decir, con menor varianza). Todo este análisis sugiere que se debe verificar la presencia del problema de la simultaneidad antes de descartar los MCO en favor de las alternativas.

El problema de la simultaneidad surge porque algunas de las regresoras son endógenas y, por consiguiente, es probable que estén correlacionadas con el término de perturbación o de error. De esta forma, en una prueba de simultaneidad, se intenta averiguar si una regresora (una endógena) está correlacionada con el término de error. Si lo está, existe el problema de simultaneidad, en cuyo caso deben encontrarse alternativas a MCO; si no lo está, se pueden utilizar MCO. Para averiguar cuál es el caso en una situación concreta, se puede utilizar la prueba del error de especificación de Hausman.

3. METODOLOGÍA

4.1 Variables y Fuente de los datos

Esta sección detalla los datos utilizados para la estimación de un modelo de predicción de la demanda eléctrica. Los datos referentes a la demanda eléctrica por tipo de usuario, número de abonados, precio de la electricidad y precio de bienes sustitutos son los disponibles en la Autoridad Reguladora de los Servicios Públicos, para el período que va desde el año 2012 hasta el mes de marzo del año 2019. Los datos relacionados a variables económicas (producción, ingreso, índice mensual de actividad económica) son tomados de las estadísticas disponibles en el sitio web del Banco Central de Costa Rica (BCCR).

El Cuadro 1 muestra la lista de variables a utilizar para la estimación de la demanda eléctrica. La información respecto a la demanda eléctrica se segmenta en tres categorías de clientes en función del grupo tarifario al que pertenecen. Para el caso de la demanda residencial se utiliza el consumo en KWh de los clientes bajo la tarifa residencial y residencial horario. La demanda del sector comercio, servicios e industriales agrupa los clientes bajo estas tarifas, incluyendo la tarifa de la categoría general. La demanda de los grandes consumidores se compone de la agregación de los abonados a la tarifa que pagan los clientes conectados a media y alta tensión.

Los precios medios de la electricidad en cada sector han sido obtenidos de manera mensual, estimados como una media aritmética simple de los precios vigentes en cada mes para cada empresa distribuidora. El número de abonados en cada sector representa el total de abonados en cada categoría de cliente.

Como variables *proxi* del ingreso se utiliza el Ingreso per-capita para el caso del sector residencial. El índice mensual de actividad económica se incluye como una variable de proxy de la producción para los demás sectores.

Tabla 1. Descripción de variables y su fuente de información

Variables	Descripción	Fuente
Qtotal	Consumo total de energía en KWh. Muestra el consumo total en un período, agregando las diferentes categorías de clientes. Es la demanda total de energía en un período dado.	ARESEP
Qres	Consumo residencial de electricidad en KWh. Comprende el consumo de electricidad de los clientes con tarifa residencial y residencial horaria	ARESEP
Qcis	Consumo total de electricidad, en KWh, de los clientes con tarifa bajo la categoría comercios, servicios e industrial.	ARESEP
Qgc	Consumo total de electricidad, en KWh, de los clientes con tarifa de media tensión, media tensión b y usuarios directo, denominados como grandes consumidores.	ARESEP
Pme	Precio medio del KWh.	ARESEP
Pres	Precio medio del KWh para el sector residencial	ARESEP
Pcis	Precio medio del KWh para el sector comercio, industria y servicios.	ARESEP
Pgc	Precio medio del KWh para el sector grandes consumidores.	ARESEP
Pglp	Precio medio del gas licuado de petróleo.	ARESEP
Ntot	Número total de abonados para el país.	ARESEP
Nres	Número de abonados del servicio residencial.	ARESEP
Ngc	Número de abonados grandes consumidores.	ARESEP
Iper	Ingreso Nacional Pér-capita.	BCCR
IMAE	Índice Mensual de Actividad Económica.	BCCR
Di	Variable dicotómica para el mes y que representa estacionalidad.	-

Fuente: Elaboración propia

4.2 Modelo de series de tiempo

Como primer modelo se desarrolla el modelo de series de tiempo Box Jenkins con fines de pronóstico, al ser los más utilizados en proyecciones de series económicas. Para ello se realiza un análisis univariante de las series considerando la demanda total y por sectores de consumo (residencial y grandes consumidores). En cada caso se procede a identificar los componentes de la serie identificando la existencia de componentes cíclicos, estacionales y de tendencia.

Posterior a lo anterior, se desarrolla una determinación de los procesos autorregresivos, de media móvil y de integración. Adicionalmente se identifican la existencia de procesos estacionales. A partir del análisis anterior se selecciona el modelo adecuado basándose en las funciones de autocorrelación simple y autocorrelación parcial. Los modelos de series de tiempo a estimar bajo la metodología Box Jenkins son los siguientes.

Estimación de la demanda eléctrica para el total país:

Un modelo del patrón estacional-no estacional lo proporciona el modelo multiplicativo sugerido por Box-Jenkins. Lo anterior supone que la serie \widehat{Qtotal}_t se le ha ajustado un modelo ARIMA(P,D,Q):

$$\begin{aligned} (1 - \Phi_s B^s - \Phi_{2s} B^{2s} - \dots - \Phi_{Ps} B^{Ps})(1 - B_s)^D \widehat{Qtotal}_t \\ = (1 - \Theta_s B^s - \Theta_{2s} B^{2s} - \dots - \Theta_{Qs} B^{Qs}) \end{aligned}$$

El primer término del lado izquierdo de la igualdad es un polinomio en B^s de grado P, llamado el operador estacional autorregresivo, el cual puede abreviarse como $\Phi_P(B^s)$. El primer término del lado derecho de la igualdad es un polinomio B^s en grado Q, llamado operador estacional de promedios móviles, que se abrevia como $\Theta_Q(B^s)$. Si se representa $(1 - B_s)^D$ mediante el símbolo ∇_s^D , el modelo estacional para \widehat{Qtotal}_t se puede escribir en forma abreviada: $\Phi_P(B^s) \nabla_s^D \widehat{Qtotal}_t = \Theta_Q(B^s) b_t$.

Si la serie contiene solo un patrón estacional, el término b_t es ruido blanco. Sin embargo, si la serie posee también un patrón no-estacional, el término b_t debe escribirse mediante un modelo no-estacional ARIMA: $\varphi_p(B) \nabla^d b_t = \theta_q(B) a_t$, donde a_t es ruido blanco. Despejando b_t se obtiene: $b_t = [\varphi_p(B) \nabla^d]^{-1} \theta_q(B) a_t$. Sustituyendo el valor de b_t en el modelo estacional se tiene:

$$\varphi_p(B) \Phi_P(B^s) \nabla^d \nabla_s^D \widehat{Qtotal}_t = \Theta_Q(B^s) \theta_q(B) a_t$$

El modelo agrupa los componentes estacionales y no-estacionales. Donde $\varphi_p(B)$ es el operador AR no estacional, $\theta_q(B)$ es el operador MA-no estacional y ∇^d es el operador de diferencia no estacional. Lo anterior se describe como un modelo ARIMA(p,d,q)(P,D,Q).

Estimación de la demanda eléctrica residencial:

$$\varphi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D\widetilde{Qres}_t = \Theta_Q(B^s)\theta_q(B)a_t$$

Estimación de la demanda eléctrica para los grandes consumidores:

$$\varphi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D\widetilde{Qgc}_t = \Theta_Q(B^s)\theta_q(B)a_t$$

4.3 Modelo de Vectores Autorregresivos

La estimación de un modelo de vectores autorregresivos para la demanda eléctrica del total país se define en función de las variables precio medio de la electricidad, el índice mensual de actividad económica, el precio de un bien sustituto como el gas licuado de petróleo, el número de abonados y componente estacional a identificar. Para los modelos VAR es posible incluir variables exógenas como dummies temporales o con tendencia en el tiempo, para este caso se considera la variable Di que incorpora el componente de estacionalidad en el modelo. El sistema de ecuaciones a utilizar es el siguiente:

$$Qtot_t = \alpha_{10} + \alpha_{11}Qtot_{t-k} + \alpha_{12}Pme_{t-k} + \alpha_{13}IMAE_{t-k} + \alpha_{14}Ntot_{t-k} + \alpha_{15}Pglp_{t-k} + \alpha_{16}D_i + \varepsilon_{1t}$$

$$Pme_t = \alpha_{20} + \alpha_{21}Qtot_{t-k} + \alpha_{22}Pme_{t-k} + \alpha_{23}IMAE_{t-k} + \alpha_{24}Ntot_{t-k} + \alpha_{25}Pglp_{t-k} + \varepsilon_{2t}$$

$$IMAE_t = \alpha_{30} + \alpha_{31}Qtot_{t-k} + \alpha_{32}Pme_{t-k} + \alpha_{33}IMAE_{t-k} + \alpha_{34}Ntot_{t-k} + \alpha_{35}Pglp_{t-k} + \varepsilon_{3t}$$

$$Ntot_t = \alpha_{40} + \alpha_{41}Qtot_{t-k} + \alpha_{42}Pme_{t-k} + \alpha_{43}IMAE_{t-k} + \alpha_{44}Ntot_{t-k} + \alpha_{45}Pglp_{t-k} + \varepsilon_{4t}$$

$$Pglp_t = \alpha_{50} + \alpha_{51}Qtot_{t-k} + \alpha_{52}Pme_{t-k} + \alpha_{53}IMAE_{t-k} + \alpha_{54}Ntot_{t-k} + \alpha_{55}Pglp_{t-k} + \varepsilon_{5t}$$

La estimación de un modelo de vectores autorregresivos para la demanda eléctrica residencial del país se define en función de las variables precio medio de la electricidad del sector residencial, el ingreso per-capita, el precio de un bien sustituto como el gas licuado de petróleo, el número de abonados y componente estacional a identificar. El sistema de ecuaciones es el siguiente:

$$Qres_t = \alpha_{10} + \alpha_{11}Qres_{t-k} + \alpha_{12}Pres_{t-k} + \alpha_{13}Iper_{t-k} + \alpha_{14}Nres_{t-k} + \alpha_{15}Pglp_{t-k} + \alpha_{16}D_i + \varepsilon_{1t}$$

$$Pres_t = \alpha_{20} + \alpha_{21}Qres_{t-k} + \alpha_{22}Pres_{t-k} + \alpha_{23}Iper_{t-k} + \alpha_{24}Nres_{t-k} + \alpha_{25}Pglp_{t-k} + \varepsilon_{2t}$$

$$Iper_t = \alpha_{30} + \alpha_{31}Qres_{t-k} + \alpha_{32}Pres_{t-k} + \alpha_{33}Iper_{t-k} + \alpha_{34}Nres_{t-k} + \alpha_{35}Pglp_{t-k} + \varepsilon_{3t}$$

$$Nres_t = \alpha_{40} + \alpha_{41}Qres_{t-k} + \alpha_{42}Pres_{t-k} + \alpha_{43}Iper_{t-k} + \alpha_{44}Nres_{t-k} + \alpha_{45}Pglp_{t-k} + \varepsilon_{4t}$$

$$Pglp_t = \alpha_{50} + \alpha_{51}Qres_{t-k} + \alpha_{52}Pres_{t-k} + \alpha_{53}Iper_{t-k} + \alpha_{54}Nres_{t-k} + \alpha_{55}Pglp_{t-k} + \varepsilon_{5t}$$

Finalmente, el modelo VAR para el caso de la demanda eléctrica de los grandes consumidores se define en función de las variables precio medio de la electricidad del sector, el índice mensual de actividad económica, el precio de un bien sustituto como el gas licuado de petróleo, el número de abonados y componente estacional a identificar. El sistema de ecuaciones es el siguiente:

$$Qgc_t = \alpha_{10} + \alpha_{11}Qgc_{t-k} + \alpha_{12}Pgc_{t-k} + \alpha_{13}IMAE_{t-k} + \alpha_{14}Ngc_{t-k} + \alpha_{15}Pglp_{t-k} + \alpha_{16}D_i + \varepsilon_{1t}$$

$$Pgc_t = \alpha_{20} + \alpha_{21}Qgc_{t-k} + \alpha_{22}Pgc_{t-k} + \alpha_{23}IMAE_{t-k} + \alpha_{24}Ngc_{t-k} + \alpha_{25}Pglp_{t-k} + \varepsilon_{2t}$$

$$\begin{aligned}
IMAE_t &= \alpha_{30} + \alpha_{31}Qgc_{t-k} + \alpha_{32}Pgc_{t-k} + \alpha_{33}IMAE_{t-k} + \alpha_{34}Ngc_{t-k} + \alpha_{35}Pglp_{t-k} \\
&\quad + \varepsilon_{3t} \\
Ncis_t &= \alpha_{40} + \alpha_{41}Qgc_{t-k} + \alpha_{42}Pgc_{t-k} + \alpha_{43}IMAE_{t-k} + \alpha_{44}Ngc_{t-k} + \alpha_{45}Pglp_{t-k} \\
&\quad + \varepsilon_{4t} \\
Pglp_t &= \alpha_{50} + \alpha_{51}Qgc_{t-k} + \alpha_{52}Pgc_{t-k} + \alpha_{53}IMAE_{t-k} + \alpha_{54}Ngc_{t-k} + \alpha_{55}Pglp_{t-k} \\
&\quad + \varepsilon_{5t}
\end{aligned}$$

El modelo VAR requiere especificar el número de rezagos óptimo a incluir en cada modelo, para ello se realiza una prueba que utiliza como referencia los criterios como el Criterio de Información Akaike (AIC, por sus siglas al inglés), Criterio Hannan–Quinn (HQ, por sus siglas al inglés), Criterio Schwarz (SC, por sus siglas al inglés) y el Criterio final de predicción del error (FPE, por sus siglas al inglés).

El criterio de información de Akaike. evalúa un modelo por lo cerca que tienden a estar los valores ajustados con respecto a los valores medios verdaderos, en términos de un cierto valor esperado. El criterio selecciona los modelos que minimizan la diferencia entre el logaritmo del valor máximo de la función de verosimilitud (\hat{L}) y el número de parámetros a estimar en el modelo (k) (Agresti, 2013):

$$AIC = -2(\ln(\hat{L}) - k)$$

El criterio de información Hannan-Quinn, realiza una penalización o ajuste por el tamaño de muestra donde $L[\hat{\theta}(k)]$ representa la función de verosimilitud de las observaciones, $\hat{\theta}(k)$ es la estimación de verosimilitud del vector de parámetros θ , n es el tamaño muestral y k el número de parámetros a estimar del modelo (Hannan y Quinn, 1979)

$$AIC = -2(\ln \hat{L}[\hat{\theta}(k)] + 2 \ln(\ln n)k)$$

El criterio final de predcción, propone elegir el conjunto de modelos que consigue el valor mínimo de la expresión:

$$\frac{1 + d/N}{1 - d/N} * 1/N * \sum_{t=1}^N \epsilon^2(t, \theta)$$

que es una estimación de la varianza del error de predicción que es obtenido cuando el modelo identificado es aplicado a otro conjunto de datos distinto del usado para la identificación. Donde d es el número total de parámetros estimados, N la longitud del número de datos muestreados y ϵ la función del error entre el valor real y el proyectado (Camacho, 2006).

4.4 Selección del modelo

Para validar los resultados obtenidos de los diferentes modelos se determinan un período de entrenamiento y un período de prueba. Se utilizará la serie de demanda eléctrica total y por tipo de cliente de enero de 2012 a octubre de 2018 para el período de entrenamiento y como período de prueba los meses de noviembre a junio de a enero de 2019. Una vez estimados los modelos se realiza el pronóstico y se comparan los resultados contra el valor real de la demanda en los meses de prueba. Para tales comparaciones se emplean la raíz del error cuadrático medio (RSME) y el valor del error porcentual absoluto medio (MAPE) los cuales se especifican en las siguientes ecuaciones:

$$RMSE = \sqrt{\frac{\sum_{t=1}^N E_t^2}{N}}$$

$$MAPE = \frac{\sum_{t=1}^N \left| \frac{E_t}{Y_t} \right|}{N}$$

Donde:

$$E_t = Y_t - F_t$$

Y se sabe que F es el pronóstico de la serie Y , y N es la cantidad de observaciones

4.5 Software a utilizar

Se utiliza el software estadístico R versión 3.6.3, el cuál es software libre; además del software Eviews Versión 10. En la selección del modelo adecuado también se analizará la

capacidad de ajustar un modelo adecuado utilizando las funciones de autocorrelación y autocorrelación parcial con las funciones teóricas, este análisis se comparará con los resultados de aplicar la función `auto.arima` de R, y la aplicación de Tramo Seat de Eviews. De esta forma también se concluirá sobre el uso de estas dos aplicaciones y su consistencia técnica y teórica para estimar el mejor modelo. Los paquetes requeridos en R para los análisis son: `forecast` (Hyndman & Khandakar, 2008), `tseries` (Trapletti [aut et al., 2013]), `urca` (Pfaff, 2008), `lmtest` (Hothorn et al., 2017), `tsstudio` (Krispin, 2020), `ggplot2` (Wickham 2009), y `dygraphs` (Vanderkam et al., 2018).

El primer paso consiste en considerar mediante el software estadístico R el análisis de la serie, su tendencia, estacionalidad, condiciones de estacionariedad y transformaciones requeridas para trabajar con una serie estacionaria. Posteriormente, se realizan y analizan las funciones de autocorrelación y autocorrelación función para identificar el o los modelos ARIMA que se someten a análisis. Al modelo seleccionado se le realizan las pruebas respectivas para determinar normalidad de los residuos, autocorrelación, bondad de ajuste entre otros.

Un segundo paso consiste en comparar el modelo seleccionado mediante el procedimiento anterior con la función `auto.arima` de la librería “Forecast” en R. Esta función elige entre una combinación de modelos aquel que minimiza el valor del Criterio de Información Akaike (AIC), Criterio de Información Akaike corregido (AICc) o el criterio de información bayesiano (BIC).

Finalmente, también se compara los resultados del modelo seleccionado en el paso uno con los resultados de la función Tramo/Seat¹ (“Regresión de series de tiempo ARIMA con ruido blanco, observaciones perdidas y valores extremos” y “Extracción de señales en modelos ARIMA) de Eviews. Esta aplicación permite la modelación de series temporales y la extracción de señales. Tramo permite estimar y pronosticar modelos de regresión con errores posiblemente no estacionarios como los ARIMA y Seats estima los componentes no observados en series temporales.

¹ Desarrollado por Victor Gomez y Agustin Maravall (1996)

4.6 Metodología de Estimación

Análisis de estacionariedad

Una serie de tiempo es estacionaria cuando la media y varianza se mantienen constantes en el tiempo, la mayoría de las series económicas son no estacionarias y requieren una transformación o diferenciación de la serie. El tests de raíz unitaria se utilizará para determinar si la serie debe ser diferenciada o hacer una regresión tendencial, en este sentido se emplearán los test de Dickey-Fuller y Dickey-Fuller Aumentado y el test de Phillips-Perron.

También en esta etapa se utilizará el test de Zivot and Andrews que permite evaluar no solo si la serie tiene raíz unitaria sino también si tiene cambios estructurales a considerar.

Etapa de identificación

Se analizan las funciones de autocorrelación y autocorrelación parcial y se comparan con las funciones teóricas para determinar el patrón AR, MA, ARIMA o SARIMA a estimar.

Etapa de estimación

En la etapa de estimación se estiman los modelos que se identificaron en la etapa anterior, principalmente se verifica la significancia de los coeficientes del modelo (AR) y (MA), además de estimar si corresponden el componente estacional ya sea MA o AR. La significancia se verifica mediante el p value. Así mismo, se revisa la bondad de ajuste de la estimación mediante el R cuadrado y R cuadrado ajustado, además del estadístico Durbin - Watson.

Etapas de verificación

En esta etapa se analizan las condiciones de invertibilidad y estacionariedad del modelo, Un modelo AR es invertible cuando los pesos de las observaciones pasadas disminuyen con el tiempo. Un modelo MA es invertible si puede ser expresado en función de los rezagos de Y_t con pesos que disminuyen con el tiempo. Un modelo AR es estacionario si depende cada vez menos de los valores del pasado, no es explosivo.

Se debe verificar que los residuos sean ruido blanco, Se evalúa con el correlograma de los residuos, estos se evalúan de forma individual mediante el correlograma revisando que se encuentren dentro de los límites de confianza, así mismo, se analiza la correlación conjunta mediante la prueba Ljung-Box. Cabe indicar que se debe analizar la normalidad de los residuos y contrastar la homocedasticidad, para lo que se utiliza la prueba formal de ARCH.

4. RESULTADOS

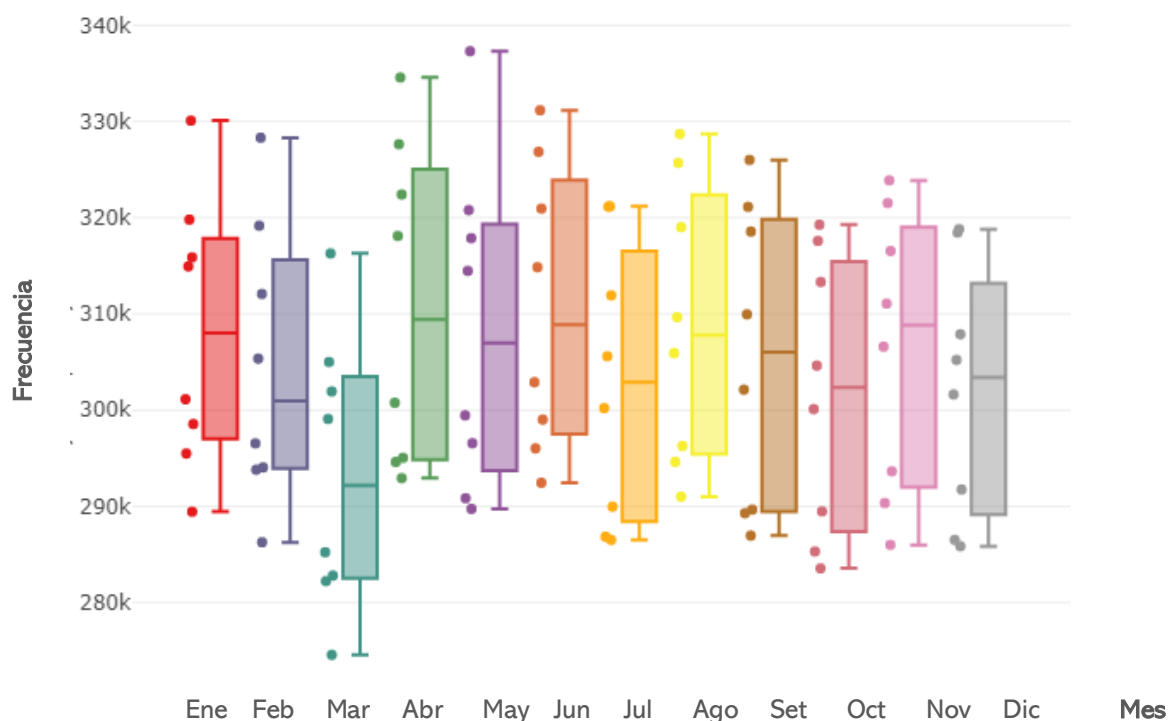
5.1 Modelos de Series de tiempo

En esta sección se presenta el análisis de la demanda eléctrica del sector residencial e industrial, para determinar su comportamiento en el periodo de análisis, identificar patrones estacionales, tendencia y valores extremos. A la vez se muestran los resultados de las pruebas aplicadas para determinar la estacionariedad o no de la serie, los modelos seleccionados y las pruebas aplicadas en cada caso.

5.1.1 Análisis de la demanda eléctrica total del sector residencial

El gráfico 1 muestra el análisis de estacionalidad para la demanda eléctrica del sector residencial del país. El consumo medio varía a lo largo del año, sin embargo, el mes de marzo muestra un consumo bajo con respecto a los demás meses del año.

Gráfico 1. Costa Rica: Estacionalidad para el consumo eléctrico residencial. Período 2012-2019



Fuente: Elaboración propia.

Análisis de estacionariedad

Existen dos procedimientos comunes para remover la tendencia de una serie temporal, primero es posible usar primeras diferencias, o bien utilizar una regresión tendencial. Las primeras diferencias suelen ser apropiadas para series temporales integradas de orden uno - $I(1)$; mientras que la regresión tendencial es apropiada para series con estacionariedad tendencial $I(0)$. Los tests de raíces unitarias pueden ser usados para determinar si una serie con tendencia primero deber ser diferenciada o hacer una regresión tendencial, esto es, si hay estacionariedad tendencial o diferencial. Por esta razón se realiza un análisis de estacionariedad a la serie de consumo eléctrico residencial.

El análisis de series de tiempo requiere que la serie sea estacionaria es decir que la media y la varianza no cambien en el tiempo. La hipótesis nula para probar es que la serie tiene raíces unitarias, por tanto, no es estacionaria. Por ende, la hipótesis alternativa es que la serie es estacionaria y k , corresponde al orden de lags o retrasos, en este caso el número de rezagos fue de 4.

El Test de Dickey-Fuller Aumentado, mostró que la serie tiene una raíz unitaria, por tanto, es no estacionaria. Se realizó una diferenciación de la serie y se aplicaron tanto el test de Dickey Fuller Aumentado y el test de Phillips- Perron, en ambos casos se rechazó la hipótesis nula de no estacionariedad, es decir la serie es estacionaria integrada de orden 1.

Cuadro 1. Resultados de las pruebas de estacionariedad para la demanda eléctrica residencial

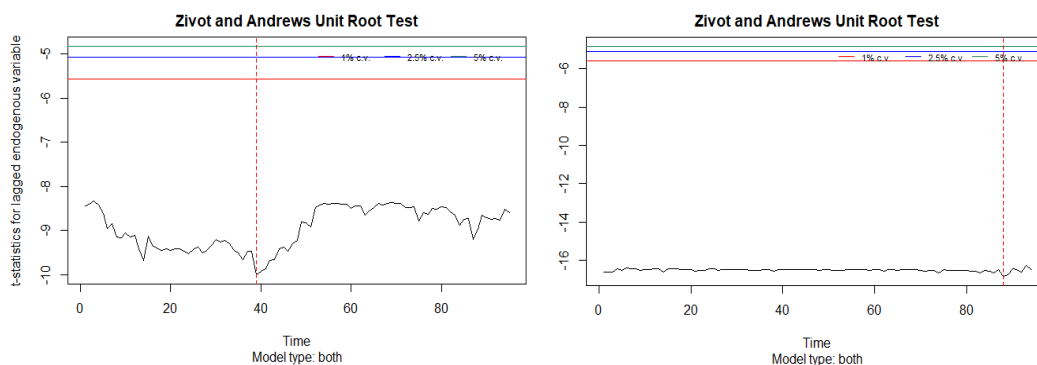
Prueba	Valor-p
Dickey-Fuller Aumentado (Serie original)	0.3343
Dickey-Fuller Aumentado (Serie en diferencias)	0.0010
Phillips- Perron (Serie en diferencias)	0.0100

Fuente: Elaboración propia

Se aplicó el test de Zivot and Andrews, como se observa en la figura 3, los resultados sugieren un cambio estructural al nivel de la observación 40, y se rechaza la hipótesis nula de raíz unitaria tal como lo sugiere el gráfico de la izquierda, el segundo gráfico muestra los

resultados con la serie diferenciada, como se observa se elimina la tendencia y variabilidad a lo largo del tiempo.

Figura 3. Resultados del Test Zivot and Andrews para determinar la presencia de raíz unitaria.

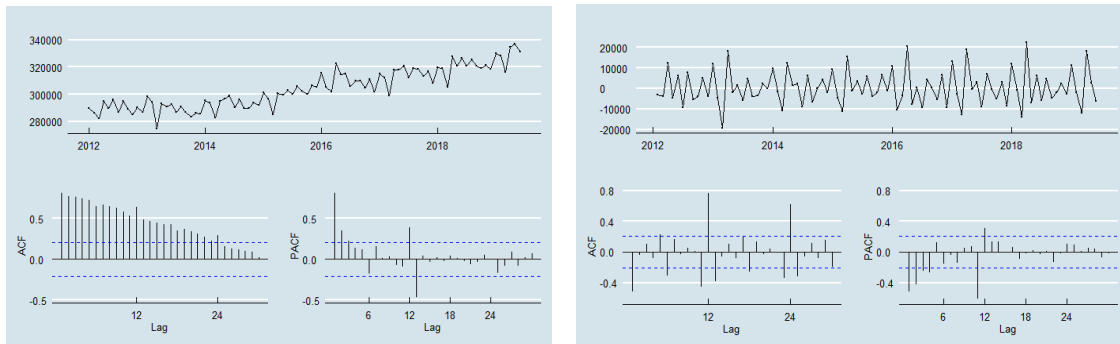


Fuente: Elaboración propia.

Etapas de identificación y estimación

Un análisis de las funciones de autocorrelación simple y parciales del consumo de energía eléctrica muestra que las correlaciones no decaen a 0 de forma rápida, tal como se muestra en el correlograma de la izquierda de la figura 4 para la serie original, con ello se considera que la serie es no estacionaria o con raíz unitaria. La corrección de los problemas de no estacionariedad y de variabilidad de la serie en el tiempo se pueden corregir mediante la aplicación de transformaciones a la serie original y la diferenciación de la serie, en el lado derecho de la figura 4 se presenta el correlograma para la serie diferenciada.

Figura 4. Funciones de autocorrelación simple y parcial para la serie original y diferenciada.



Fuente: Elaboración propia.

La función de autocorrelación (ACF, siglas en inglés) mide la correlación entre dos variables separadas por k periodos. Mide el grado de asociación lineal que existe entre dos variables del mismo proceso estocástico. La función de autocorrelación parcial (PACF, siglas en inglés) mide la correlación entre dos variables separadas por k periodos cuando no se considera la dependencia creada por los retardos intermedios existentes entre ambas, mide la autocorrelación que existe entre dos variables separadas k períodos descontando los posibles efectos debidos a variables intermedias.

Generalmente, la función ACF es usada para identificar el proceso de media móvil (MA, siglas al inglés) en un modelo ARIMA; mientras que la función PACF se usa para identificar los valores de la parte del proceso autoregresivo (AR). Los gráficos generados a partir de estas funciones se denominan correlogramas, los cuales contienen bandas con un intervalo de confianza del 95%.

Con el objetivo de identificar un patrón asociado a las funciones teóricas de autocorrelación simple y parcial, seguidamente se analizan estas funciones para la serie estacionaria (figura 2, correlograma derecho). Del análisis del correlograma se observan varios probables procesos, principalmente es claro un proceso MA(1) que se identifica mediante la observación de la función de autocorrelación y su comparación con las funciones teóricas de autocorrelación., por su parte, también se observan en la función de autocorrelación parcial

procesos de orden MA, lo anterior en la parte estacional. En la parte estacional del correlograma se observa que no existe un decaimiento a 0, con lo cual se concluye que es necesario diferenciar la serie en el componente estacional.

Los resultados del modelo seleccionado se presentan en el Cuadro 2. El modelo presenta un ajuste de 0.49 medido mediante el coeficiente de bondad de ajuste (R), el error de regresión es de 0.011 mientras que el valor asociado al Durbin y Watson es de 2.0114. Este es el modelo seleccionado para continuar con la etapa de verificación.

Cuadro 2. Resultados del modelo ajustado: ARIMA (0,1,1) (0,1,1)₁₂

Variable	Coeficiente	Error estándar	Valor - z	Prob.
MA(1)	-0.5085	0.1031	-4.9340	8.057e-07 ***
SMA(1)	-0.9985	0.2827	-3.5325	8.057e-07 ***

Fuente: Elaboración propia

Etapas de verificación

Condiciones de Invertibilidad y estacionariedad

Un modelo MA siempre cumple la condición de estacionariedad tanto en la parte regular como en la parte estacional. Lo anterior se comprueba dado que las raíces características de la ecuación están fuera del círculo unitario. *El modelo* si es invertible, dado que los coeficientes asociados a MA(1) y SMA(1) son menores a la unidad.

Significancia de los parámetros

La significancia en los parámetros se verifica mediante una prueba de hipótesis para el contraste de los coeficientes asociados a MA y SMA, en cada caso se prueba la hipótesis nula de que cada coeficiente es igual a cero. Por tanto, el contraste a considerar es $H_0: \theta_1 = 0$. Dados los resultados del modelo presentado en el cuadro 3, existe suficiente evidencia estadística para rechazar la hipótesis nula de que los coeficientes son estadísticamente iguales

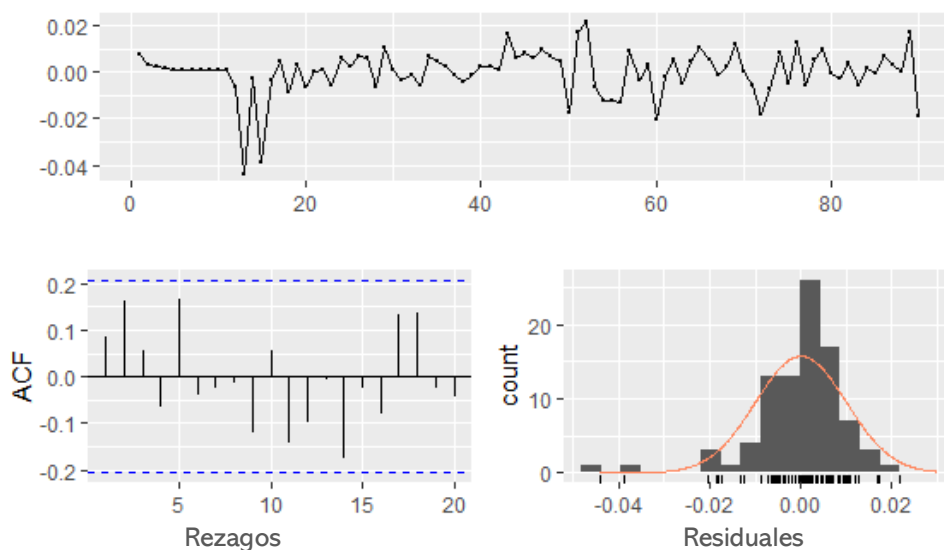
a 0. Es decir, hay evidencia esta los coeficientes son estadísticamente significativos dado que el valor “p” asociado a cada uno de ellos es menor al nivel de significancia del 5%.

Pruebas de residuos

Con los resultados del correlograma que se incluye en la figura 5, se verifica la significancia de las correlaciones individuales, es decir, para cada rezago (k) se debe probar $H_0: \rho_k = 0$. Las correlaciones muestrales r_k se encuentran dentro de los límites de confianza, con lo que no se rechaza la H_0 .

Es decir, no existe evidencia estadística para rechazar la hipótesis nula de que los residuos no están correlacionados, es decir los residuos provenientes del modelo ARIMA no están correlacionados considerando hasta 10 rezagos. Dado que el p valor en la prueba Ljung-Box arroja un valor de 0.3781, con lo cual se asume que el modelo no tiene correlación serial y es adecuado utilizar este modelo.

Figura 5. Residuos del modelo seleccionado (0,1,1)(0,1,1)[12], función de autocorrelación e histograma.



Fuente: Elaboración propia.

Pruebas para la media de los residuos

Se prueba la hipótesis de que la media de los residuos es igual a 0. Por tanto, la hipótesis nula se establece como $H_0: \mu_{(at)} = 0$. De acuerdo con los resultados no se rechaza la H_0 de media 0.

Análisis de varianza y homocedasticidad

El supuesto de correlaciones de la varianza de los residuos se analiza con el correlograma de los residuos al cuadrado (Ver anexo 5). Como se observa solo el rezago $K=26$ se sale de las bandas, por lo que es un indicio para ese rezago de algún problema de heteroscedasticidad. El supuesto que la varianza no está correlacionada (de manera conjunta) también se revisa en el correlograma de los residuos al cuadrado. Como se aprecia, para los primeros residuos (k) la probabilidad es mayor al 5% de significancia, por lo que no se rechaza la Hipótesis nula conjunta de que no existen correlaciones de la varianza de los residuos.

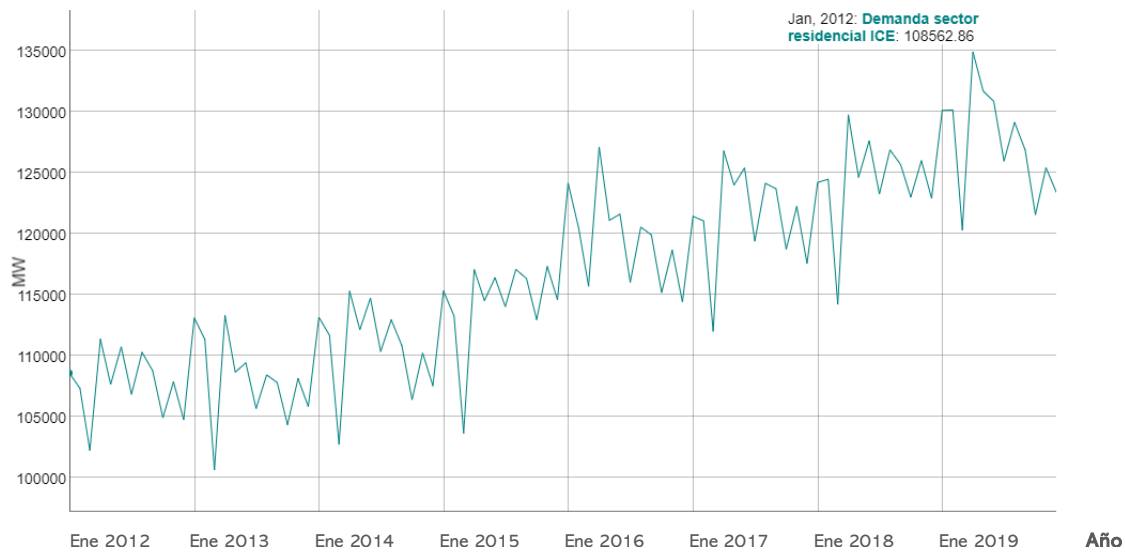
Prueba de homocedasticidad

Para contrastar la homocedasticidad de los residuos se realiza la prueba formal de ARCH. La H_0 que se prueba es que los residuos son homocedásticos. De los resultados se desprende que la probabilidad “p” asociada a la prueba de homocedasticidad de ARCH (0.83) es mayor al nivel de significancia del 5%, por tanto, no se rechaza la H_0 de homocedasticidad de los residuos, es decir se asume que los residuos presentan varianza constante.

5.1.2 Análisis de la demanda eléctrica total del sector residencial para el ICE

El gráfico 2 muestra el comportamiento de la demanda eléctrica residencial para el Instituto Costarricense de Electricidad, se observa una tendencia creciente durante todo el período, sin embargo, con una tendencia a la baja durante el segundo trimestre del 2019.

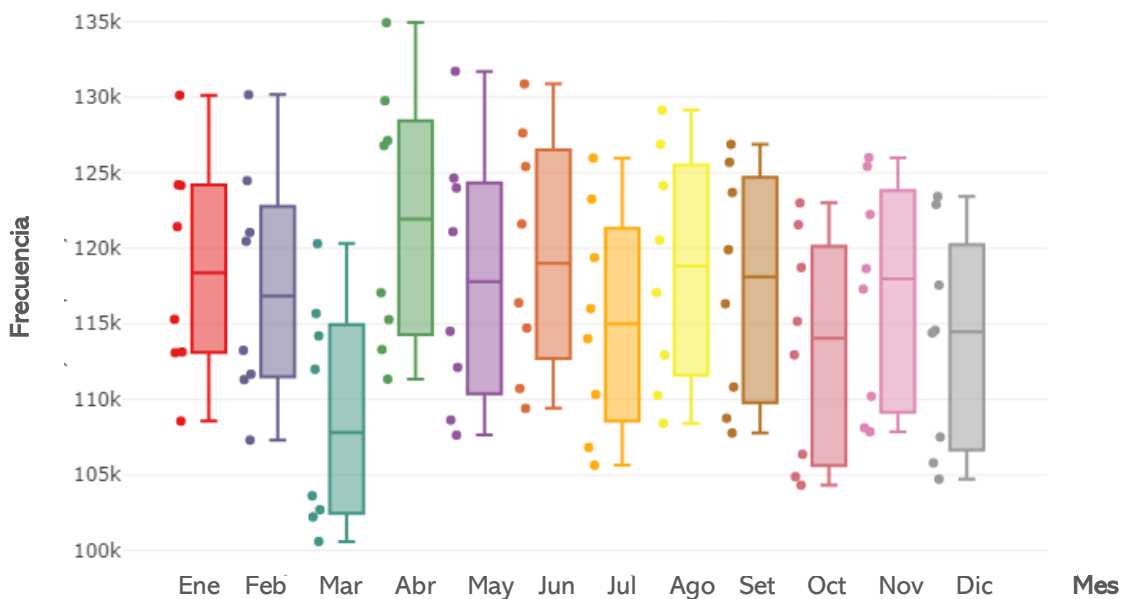
Gráfico 2. ICE: Evolución de la demanda eléctrica residencial. Período 2012-2019.



Fuente: Elaboración propia.

Mediante el gráfico 3 se analiza la estacionalidad de la demanda eléctrica del sector residencial para el ICE, su comportamiento tiende ser similar a la demanda nacional, sin embargo, se observa que la media varía un poco más a lo largo del año. No se encuentra un patrón marcado de estacionalidad, es decir, todos los meses la medida de las ventas varía, sin embargo, los meses de marzo, julio, octubre y diciembre presentan una demanda media menor en comparación con los otros meses del año.

Gráfico 3. ICE: Estacionalidad de la demanda eléctrica residencial. Período 2012-2019.



Fuente: Elaboración propia.

Análisis de estacionariedad

El Test de Dickey-Fuller Aumentado, mostró que la serie tiene una raíz unitaria, por tanto, es no estacionaria. Se realizó una diferenciación de la serie y se aplicaron los test de Dickey Fuller Aumentado y el test de Phillips- Perron, como se observa en el Cuadro 3 en ambos casos se rechazó la hipótesis nula de no estacionariedad. Es decir, la serie es estacionaria integrada de orden 1.

Cuadro 3. Resultados de las pruebas de estacionariedad para la demanda eléctrica residencial.

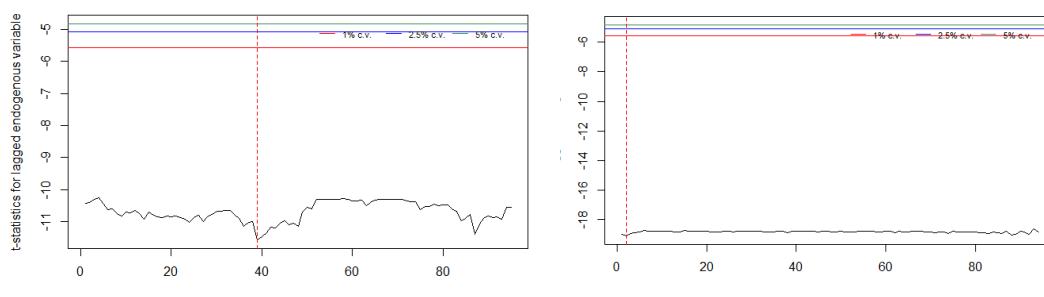
Prueba	Valor-p
Dickey-Fuller Aumentado (Serie original)	0.2911
Dickey-Fuller Aumentado (Serie en diferencias)	0.0010
Phillips- Perron (Serie en diferencias)	0.0100

Fuente: Elaboración propia.

Se aplicó el test de Zivot and Andrews, como se observa en la figura 6, al igual que sucede para la demanda nacional, los resultados sugieren un cambio estructural al nivel de la

observación 40, y se rechaza la hipótesis nula de raíz unitaria, el segundo gráfico muestra los resultados con la serie diferenciada, se observa claramente que la misma se estabiliza.

Figura 6. Resultados del Test Zivot and Andrews para determinar la presencia de raíz unitaria.

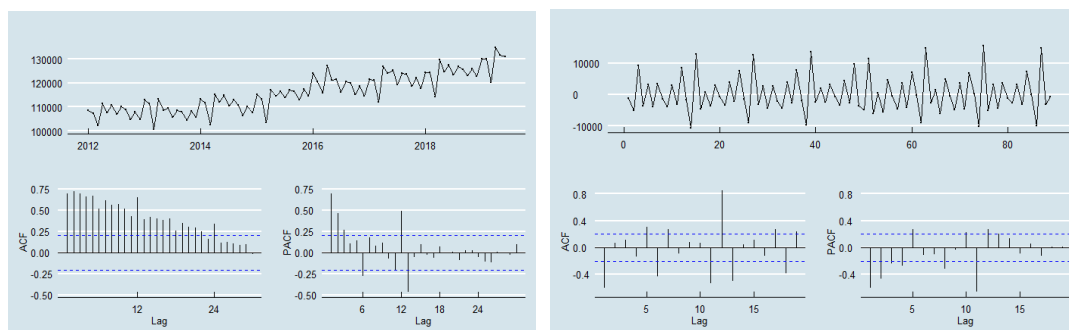


Fuente: Elaboración propia.

Etapas de identificación y estimación

Las funciones de autocorrelación simple y parciales del consumo de energía eléctrica muestra que las correlaciones no decaen a 0 de forma rápida, con ello se considera que la serie es no estacionaria o con raíz unitaria. La corrección de los problemas de no estacionariedad y de variabilidad de la serie en el tiempo se pueden corregir mediante la aplicación de transformaciones a la serie original y la diferenciación de la serie, en el lado derecho de la figura 7 se presenta el correlograma para la serie diferenciada.

Figura 7. Funciones de autocorrelación simple y parcial para la serie original y diferenciada.



Fuente: Elaboración propia.

Las funciones teóricas de autocorrelación simple y parcial se comparan con las de la figura 7, con el objetivo de identificar un patrón ARIMA. Del análisis del correlograma se observan varios probables procesos, en la parte no estacional se observa un proceso MA(1) a partir del análisis de las funciones de autocorrelación de la serie diferenciada. Adicionalmente, se observa algún componente estacional. Se verifico inicialmente un modelo SARIMA(0,1,1)(0,1,1)₁₂ sin embargo presentó problemas de autocorrelación y no normalidad de los residuos.

Los resultados del modelo seleccionado se presentan en el Cuadro 4. Este es el modelo seleccionado para continuar con la etapa de verificación.

Cuadro 4. Resultados del modelo ajustado: SARIMA (0,1,1) (1,1,0)₁₂

Variable	Coeficiente	Error estándar	Valor - z	Prob.
MA(1)	-0.2526	0.1066	-2.3700	0.0178 *
SAR(1)	-0.5015	0.1039	-4.8258	1.394e-06 ***

Fuente: Elaboración propia

Etapas de verificación

Condiciones de Invertibilidad y estacionariedad

Un modelo MA siempre cumple la condición de estacionariedad tanto en la parte regular como en la parte estacional. Lo anterior se comprueba dado que las raíces características de la ecuación están fuera del círculo unitario. *El modelo* si es invertible, dado que los coeficientes asociados a MA(1) y SAR(1) son menores a la unidad.

Significancia de los parámetros

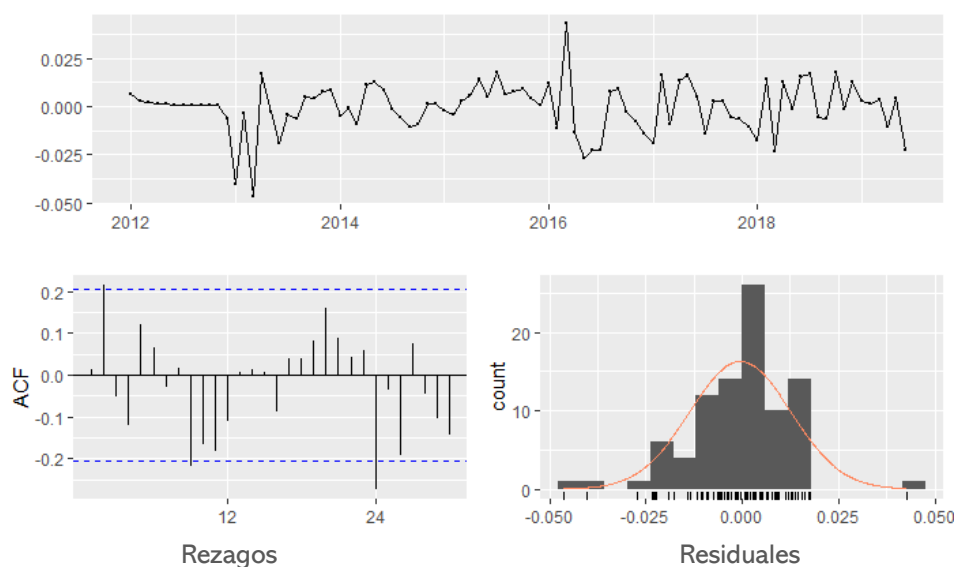
La significancia en los parámetros se verifica mediante una prueba de hipótesis para el contraste de los coeficientes asociados a MA y SAR, en cada caso se prueba la hipótesis nula de que cada coeficiente es igual a cero. Hay evidencia esta los coeficientes son

estadísticamente significativos dado que el valor “p” asociado a cada uno de ellos es menor al nivel de significancia del 5%.

Pruebas de residuos

Con los resultados del correlograma que se incluye en la figura 8, se verifica la significancia de las correlaciones individuales. Mediante la función de autocorrelación parcial se observa un solo rezago estadísticamente distinto de 0. Se realiza la prueba formal de Ljung-Box, considerando la hipótesis nula de que los residuos están independientemente distribuidos, con un valor $p=0,15$ no existe evidencia estadística suficiente para rechazar la hipótesis nula de no autocorrelación serial.

Figura 8. Residuos del modelo seleccionado $(0,1,1)(1,1,0)[12]$, función de autocorrelación e histograma.



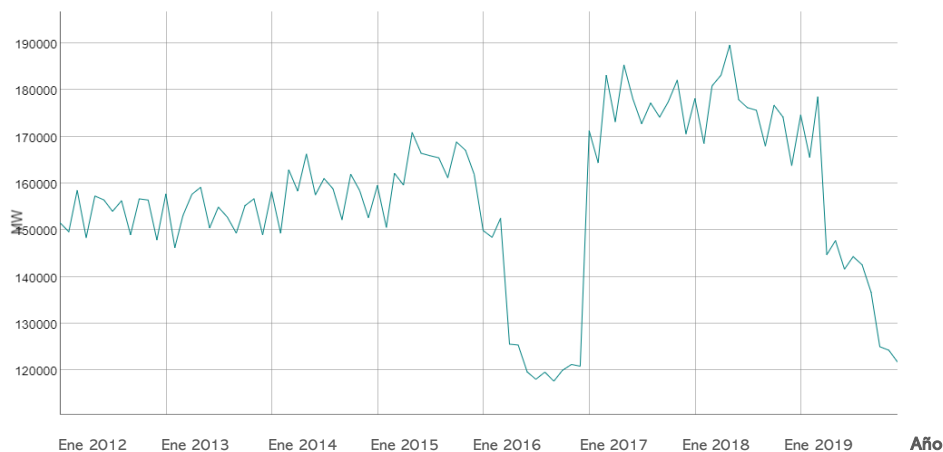
Fuente: Elaboración propia.

5.1.3 Análisis de la demanda eléctrica total para grandes consumidores

El gráfico 4 se presenta el comportamiento de la demanda eléctrica para el sector industrial, que ha sido catalogado como el sector de grandes consumidores (Tarifa Tm-a). En este caso es el consumo o demanda eléctrica para el total país. Se observa durante el período 2012-

2014 un comportamiento muy similar sin una tendencia clara, sin embargo, a partir del 2016 la tendencia en la demanda eléctrica cae significativamente y se recupera a inicios del 2017. Par el año 2019 muestra una tendencia a disminuir el consumo eléctrico de este sector.

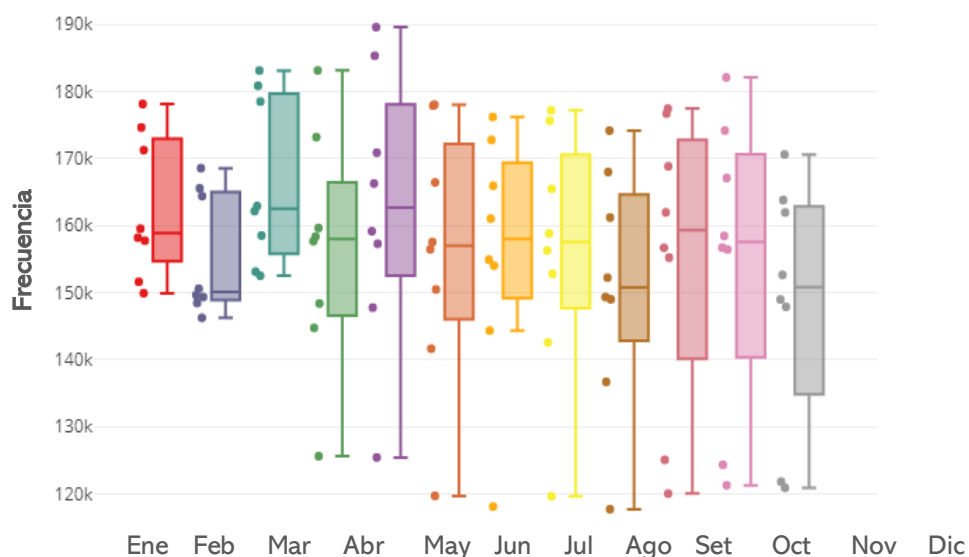
Gráfico 4. Costa Rica: Evolución de la demanda eléctrica para grandes consumidores. Período 2012-2019.



Fuente: Elaboración propia

La estacionalidad para la demanda de los grandes consumidores se analiza en el gráfico 5. Lo meses de febrero presentan una demanda media menor en comparación con los demás meses del año, y se observa en términos generales una media similar para todos los meses, sin embargo con valores extremos bajos entre abril y diciembre.

Gráfico 5. Costa Rica: Estacionalidad para el consumo eléctrico de grandes consumidores. Período 2012-2019



Fuente: Elaboración propia.

Análisis de estacionariedad

El análisis de estacionariedad de la serie se realiza mediante el Test de Dickey-Fuller Aumentado, donde se observa que la serie tiene una raíz unitaria ($p\text{-valor} = 0.4673$), por tanto, es no estacionaria. Se realizó una diferenciación de la serie y se aplicaron los test de Dickey Fuller Aumentado ($p\text{-valor} = 0.0525$) y el test de Phillips- Perron ($p\text{-valor} = 0.01$), como se observa en el Cuadro 5 en ambos casos se rechazó la hipótesis nula de no estacionariedad. Con esto se asume que la serie es estacionaria integrada de orden 1.

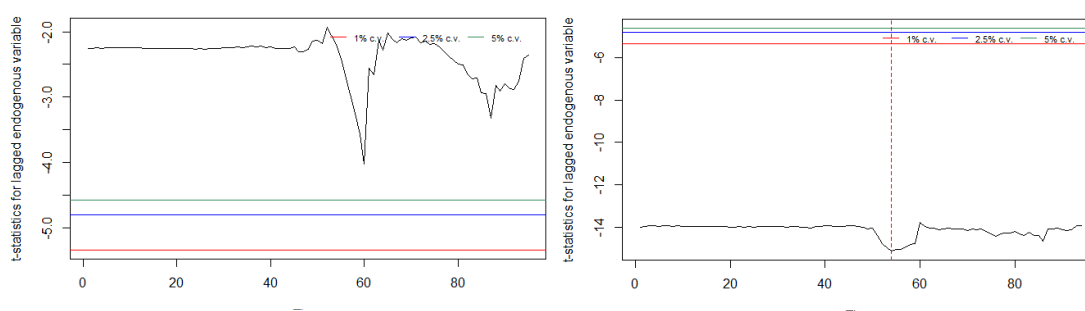
Cuadro 5. Resultados de las pruebas de estacionariedad para la demanda eléctrica grandes consumidores.

Prueba	Valor-p
Dickey-Fuller Aumentado (Serie original)	0.4673
Dickey-Fuller Aumentado (Serie en diferencias)	0.0525
Phillips- Perron (Serie en diferencias)	0.0100

Fuente: Elaboración propia.

Adicionalmente, se aplicó el test de Zivot and Andrews, como se observa en la figura 9, se comprueba de manera gráfica la necesidad de diferenciar la serie, el gráfico de la izquierda muestra que tanto la media como la varianza no son constantes en el tiempo. Los resultados sugieren un cambio estructural al nivel de la observación 54, una vez diferenciada la serie y se rechaza la hipótesis nula de raíz unitaria, el segundo gráfico muestra los resultados con la serie diferenciada.

Figura 9. Resultados del Test Zivot and Andrews para determinar la presencia de raíz unitaria.



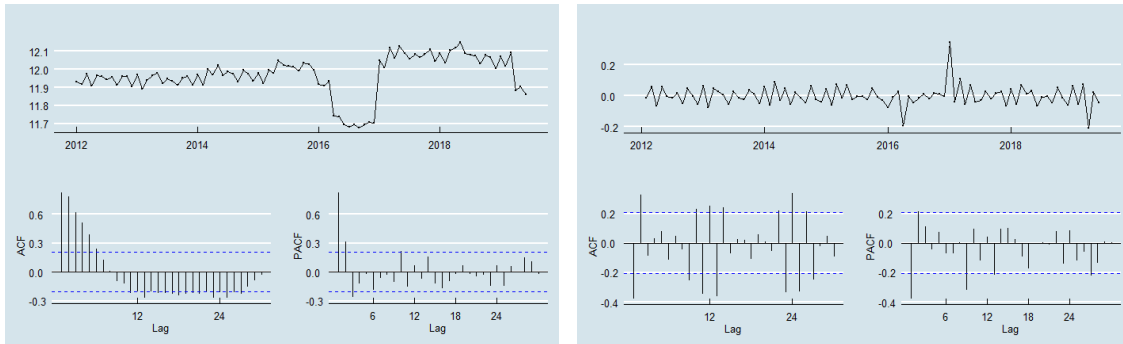
Fuente: Elaboración propia.

Etapas de identificación y estimación

Las funciones de autocorrelación simple y parciales del consumo de energía eléctrica para grandes consumidores muestra que las correlaciones no decaen a 0 de forma rápida, con ello se considera que la serie es no estacionaria o con raíz unitaria, como se identificó anteriormente. El gráfico de la derecha muestra la serie con una transformación logarítmica y diferenciada.

Del análisis del correlograma se observan varios probables procesos, en la parte no estacional se observa un proceso AR(1).

Figura 10. Funciones de autocorrelación simple y parcial para la serie original y diferenciada.



Fuente: Elaboración propia.

Analizados varios modelos, en el Cuadro 6, se presentan los resultados del modelo seleccionado utilizando el criterio AIC.

Cuadro 6. Resultados del modelo ajustado: SARIMA (1,1,1) (1,1,0)₁₂

Variable	Coficiente	Error estándar	Valor - z	Prob.
AR(1)	0.7148	0.2414	2.9605	0.0031 ***
MA(1)	-0.6042	0.2602	-2.3222	0.0202 *
SAR(1)	-0.5499	0.0895	-6.1451	7.992 e -10 ***

Fuente: Elaboración propia

Etapa de verificación

Condiciones de Invertibilidad y estacionariedad

Un modelo AR siempre cumple la condición de estacionariedad tanto en la parte regular como en la parte estacional. Lo anterior se comprueba dado que las raíces características de la ecuación están fuera del círculo unitario. El modelo si es invertible, dado que los coeficientes asociados a AR(1), MA(1) y SAR(1) son menores a la unidad.

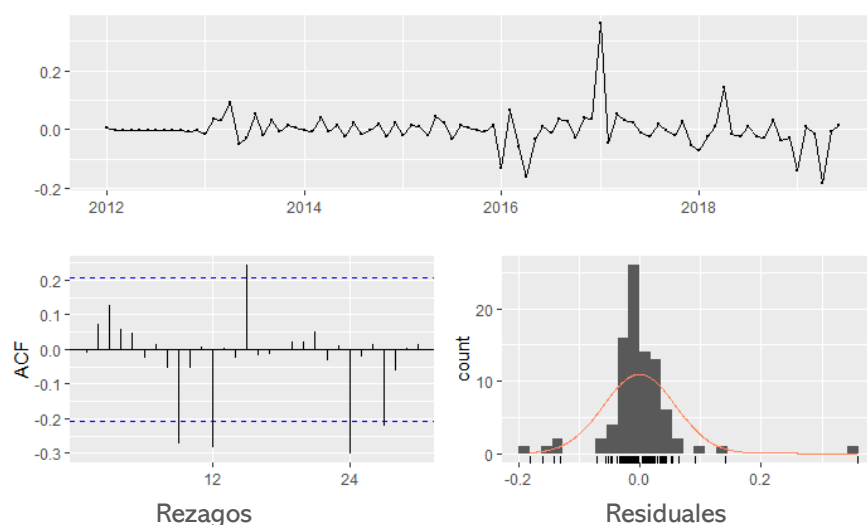
Significancia de los parámetros

La significancia en los parámetros se verifica mediante una prueba de hipótesis para el contraste de los coeficientes asociados a AR, MA y SAR, en cada caso se prueba la hipótesis nula de que cada coeficiente es igual a cero. Hay evidencia esta los coeficientes son estadísticamente significativos dado que el valor “p” asociado a cada uno de ellos es menor al nivel de significancia del 5%.

Pruebas de residuos

Con los resultados del correlograma que se incluye en la figura 11, se verifica la significancia de las correlaciones individuales. Mediante la función de autocorrelación parcial se observa cuatro rezagos estadísticamente distintos de 0. Se realiza la prueba formal de Ljung-Box, considerando la hipótesis nula de que los residuos están independientemente distribuidos, con un valor $p=0,0376$ existe evidencia estadística suficiente para rechazar la hipótesis nula de no autocorrelación serial.

Figura 11. Residuos del modelo seleccionado $(1,1,1)(1,1,0)[12]$, función de autocorrelación e histograma.

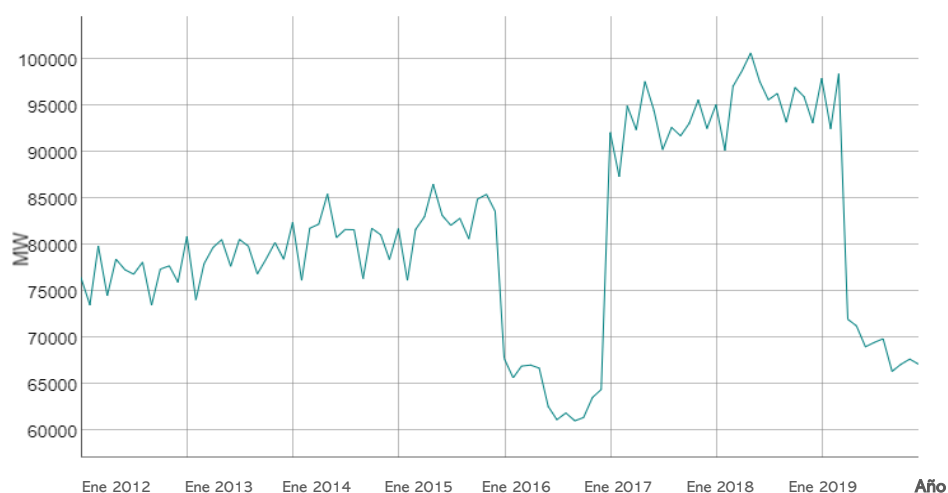


Fuente: Elaboración propia.

5.1.4 Análisis de la demanda eléctrica total para grandes consumidores del ICE

El gráfico 6 presenta el comportamiento de la demanda eléctrica para los grandes consumidores del ICE durante el período en análisis. Durante el periodo 2012-2015 el comportamiento muestra una leve tendencia al alza, sin embargo, posteriormente cae significativamente durante el año 2016 y se recupera el en 2017, volviéndose a disminuir la demanda durante el primer semestre del 2019.

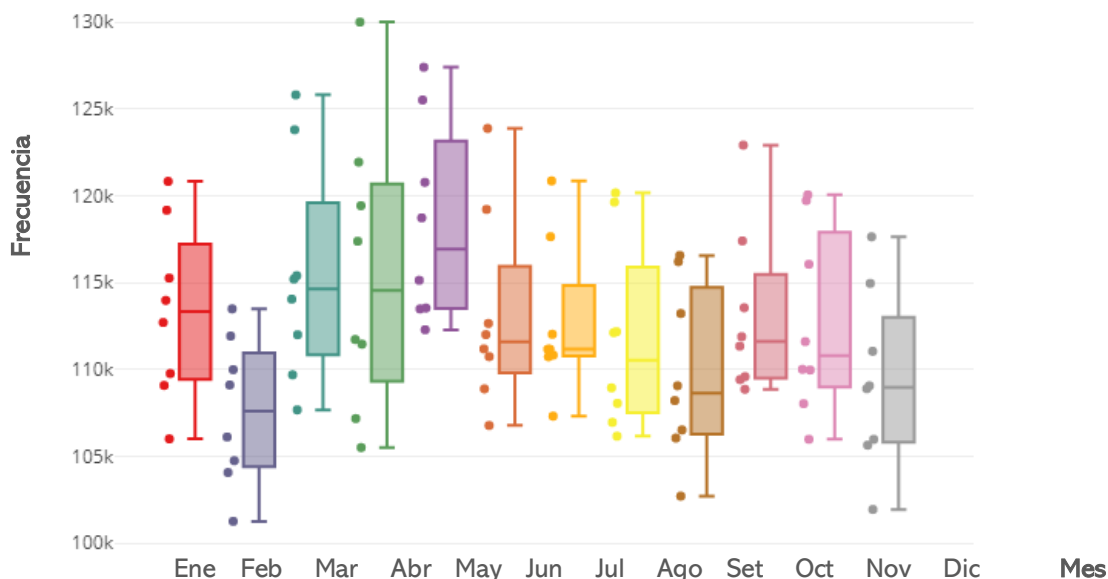
Gráfico 6. ICE: Evolución de la demanda eléctrica para grandes consumidores. Período 2012-2019.



Fuente: Elaboración propia.

En el gráfico 7, se analiza el comportamiento estacional de la demanda eléctrica. Los meses de febrero, setiembre y diciembre muestran una menor demanda media en comparación con los demás meses. Por otro lado, abril y mayo son meses con una demanda media alta en comparación al resto del año.

Gráfico 7. Costa Rica: Estacionalidad para el consumo eléctrico de grandes consumidores. Período 2012-2019.



Fuente: Elaboración propia.

Análisis de estacionariedad

Inicialmente, el análisis de estacionariedad de la serie se realiza mediante el Test de Dickey-Fuller Aumentado, donde se observa que la serie tiene una raíz unitaria ($p\text{-valor} = 0.4625$), por tanto, es no estacionaria. Se realizó una diferenciación de la serie y se aplicaron el test de Dickey Fuller Aumentado ($p\text{-valor} = 0.07842$) sin embargo este test no mostro evidencia para descartar la hipótesis nula de raíz unitaria en la serie. Sin embargo, se aplicaron dos test adicionales el test de Phillips- Perron ($p\text{-valor} = 0.01$) y el test de Kwiatkowski–Phillips–Schmidt–Shin ($p\text{-valor} = 0.01$) en ambos casos se rechazó la hipótesis nula de no estacionariedad. Con esto se asume que la serie es estacionaria integrada de orden 1 una vez diferenciada.

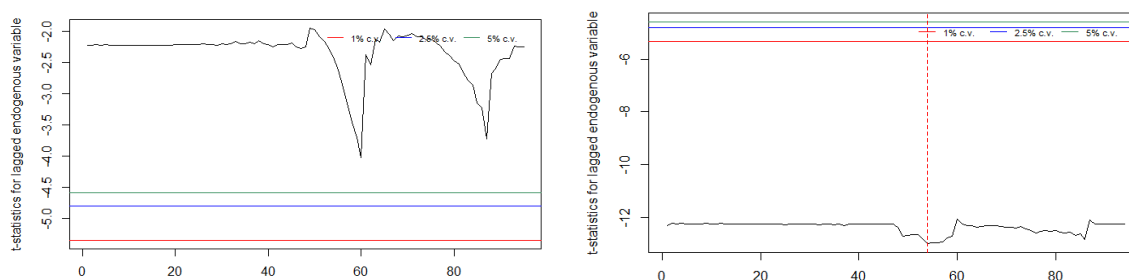
Cuadro 7. Resultados de las pruebas de estacionariedad para la demanda eléctrica grandes consumidores del ICE.

Prueba	Valor-p
Dickey-Fuller Aumentado (Serie original)	0.4625
Dickey-Fuller Aumentado (Serie en diferencias)	0.0784
Phillips- Perron (Serie en diferencias)	0.0100

Fuente: Elaboración propia.

Se aplicó el test de Zivot and Andrews, como se observa en la figura 12, se comprueba de manera gráfica la necesidad de diferenciar la serie, el gráfico de la izquierda muestra que tanto la media como la varianza no son constantes en el tiempo. Los resultados sugieren un cambio estructural al nivel de la observación 54 una vez diferenciada la serie, el segundo gráfico muestra los resultados con la serie diferenciada.

Figura 12. Resultados del Test Zivot and Andrews para determinar la presencia de raíz unitaria.



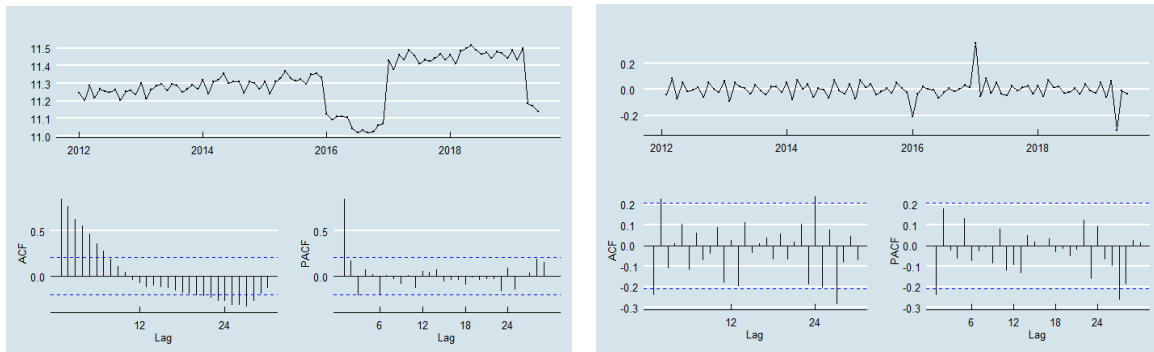
Fuente: Elaboración propia.

Etapa de identificación y estimación

Las funciones de autocorrelación simple y parciales del consumo de energía eléctrica para grandes consumidores del ICE muestra que las correlaciones no decaen a 0 de forma rápida, con ello se considera que la serie es no estacionara o con raíz unitaria, como se identificó anteriormente. EL gráfico de la derecha muestra la serie con una transformación logarítmica y diferenciada.

Del análisis del correlograma se observan varios probables procesos, en la parte no estacional inicialmente se pude considerar un proceso AR en la parte no estacional.

Figura 13. Funciones de autocorrelación simple y parcial para la serie original y diferenciada.



Fuente: Elaboración propia.

Analizados varios modelos, en el Cuadro 8, se presentan los resultados del modelo seleccionado utilizando el criterio AIC.

Cuadro 8. Resultados del modelo ajustado: SARIMA (0,1,0) (0,1,1)₁₂.

Variable	Coefficiente	Error estándar	Valor - z	Prob.
SMA(1)	-0.9999	0.2011	-4.973	6.592 e -07 ***

Fuente: Elaboración propia

Significancia de los parámetros

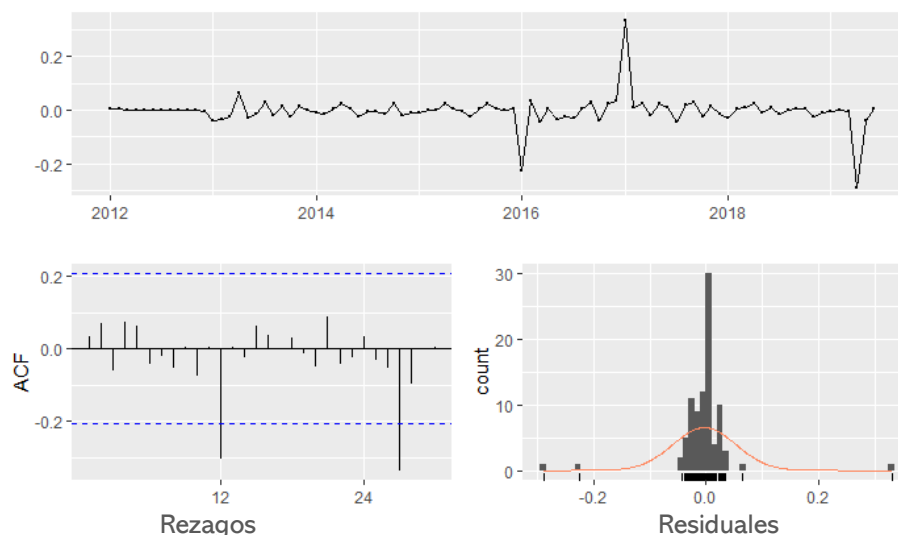
La significancia en los parámetros se verifica mediante una prueba de hipótesis para el contraste de los coeficientes asociados a SMA(1), en cada caso se prueba la hipótesis nula de que cada coeficiente es igual a cero. Hay evidencia esta los coeficientes son estadísticamente significativos dado que el valor “p” asociado a cada uno de ellos es menor al nivel de significancia del 5%.

Pruebas de residuos

La figura 14 muestra el análisis de los residuos. Mediante la función de autocorrelación parcial se observa un solo rezago estadísticamente distinto de 0. Se realiza la prueba formal de Ljung-Box, considerando la hipótesis nula de que los residuos están independientemente

distribuidos, con un valor $p=0,7033$ no existe evidencia estadística suficiente para rechazar la hipótesis nula de no autocorrelación serial.

Figura 14. Residuos del modelo seleccionado $(0,1,0)(0,1,1)[12]$, función de autocorrelación e histograma.



Fuente: Elaboración propia.

5.2 Modelos de Vectores Autorregresivos (VAR)

En la siguiente sección se presentan los resultados de la aplicación de un modelo VAR para determinar la demanda de consumo eléctrico, en función una serie de variables socioeconómicas. Tanto la demanda para grandes consumidores, como para los consumidores del sector residencial se modelo utilizando variables como el precio medio de la electricidad, el precio de un bien sustituto (gas licuado de petróleo), la cantidad de abonados, el índice mensual de actividad económica, el índice de precios al consumidor y el índice al productor industrial.

En todos los casos se realizó una prueba para determinar el número de rezagos óptimo a incluir en cada modelo utilizando criterios como el Criterio de Información Akaike (AIC, por sus siglas al inglés), Criterio Hannan–Quinn (HQ, por sus siglas al inglés), Criterio Schwarz (SC, por sus siglas al inglés) y el Criterio final de predicción del error (FPE, por sus siglas al inglés). Posteriormente, dado el número de rezagos seleccionado se procedió a realiza un

modelo con las variables seleccionadas en cada caso, y realizar un proceso “stepwise” tanto para la selección de las variables como para el número de rezagos estadísticamente distintos de cero; esto se realiza mediante el diseño de una matriz de restricciones que se incluye en la estimación de un nuevo modelo VAR.

Los modelos VAR estiman un conjunto de ecuaciones de manera simultánea, sin embargo, a continuación, se presentan en cada caso solo los resultados de la ecuación de demanda eléctrica que es la de interés. Para efectos de estimación y predicción debe tenerse en cuenta que se estiman las ecuaciones para todas las variables en su conjunto.

5.2.1 Demanda total grandes consumidores.

El cuadro 9 presenta los resultados del modelo VAR estimado para las variables: consumo eléctrico de los grandes consumidores (R_GRANDES), índice mensual de actividad económica (IMAE), precio medio de la tarifa media tensión a (PMET_TOT) y cantidad de abonados del sector (Ab_Gran_tot). Para el caso la estimación de la demanda eléctrica resultó estadísticamente significativas el consumo, el IMAE y la cantidad de abonados rezagadas un período. El grado de bondad de ajuste del modelo medido por el Coeficiente R es de 0.99.

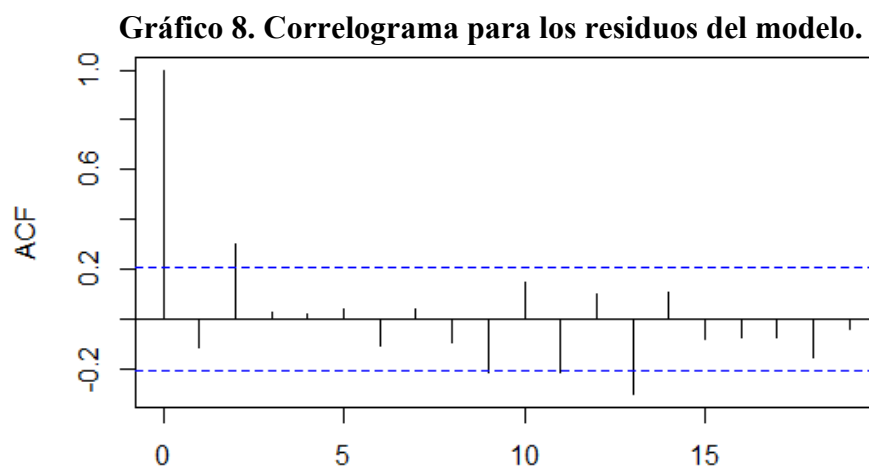
Cuadro 9. Resultados del modelo VAR para la demanda eléctrica de grandes consumidores.

	Estimación	Error estándar	Valor-t	Pr(> t)
R_GRANDES.l1	0.0674	739.0	9.120	3.09e-14 ***
IMAE.l1	-918.3	267.7	-3.430	0.000934 ***
Ab_Gran_tot.l1	302.10	80.83	3.737	0.000336 ***
Constante	-137000	4.6800	-2.928	0.004377 **

Fuente: Elaboración propia.

El modelo cumple con la condición de estabilidad o prueba de estacionariedad de múltiples variables. Dado que las raíces del polinomio se encuentran dentro del círculo unitario, es decir, son menores a 1 (0.9914, 0.8257, 0.8257, 0.0476).

La autocorrelación serial de orden superior se evaluó mediante la prueba de Portmanteau multivariada, dado un valor $p = 0,1294$; no existe evidencia estadística suficiente para rechazar la hipótesis nula de no autocorrelación serial. Igualmente, el gráfico 8 se observa que los residuos caen rápidamente a 0 sin un patrón definido evidenciando la no presencia de autocorrelación.



Fuente: Elaboración propia

Adicionalmente, se analiza la normalidad de las variables mediante la prueba de Jarque Bera (JB), a nivel individual solo la variable del IMAE no rechaza la hipótesis nula de normalidad ($p=0.9061$), a nivel global se rechaza la hipótesis nula de normalidad de los residuos del modelo ($p=2.2e-16$) evaluado mediante la prueba de JB.

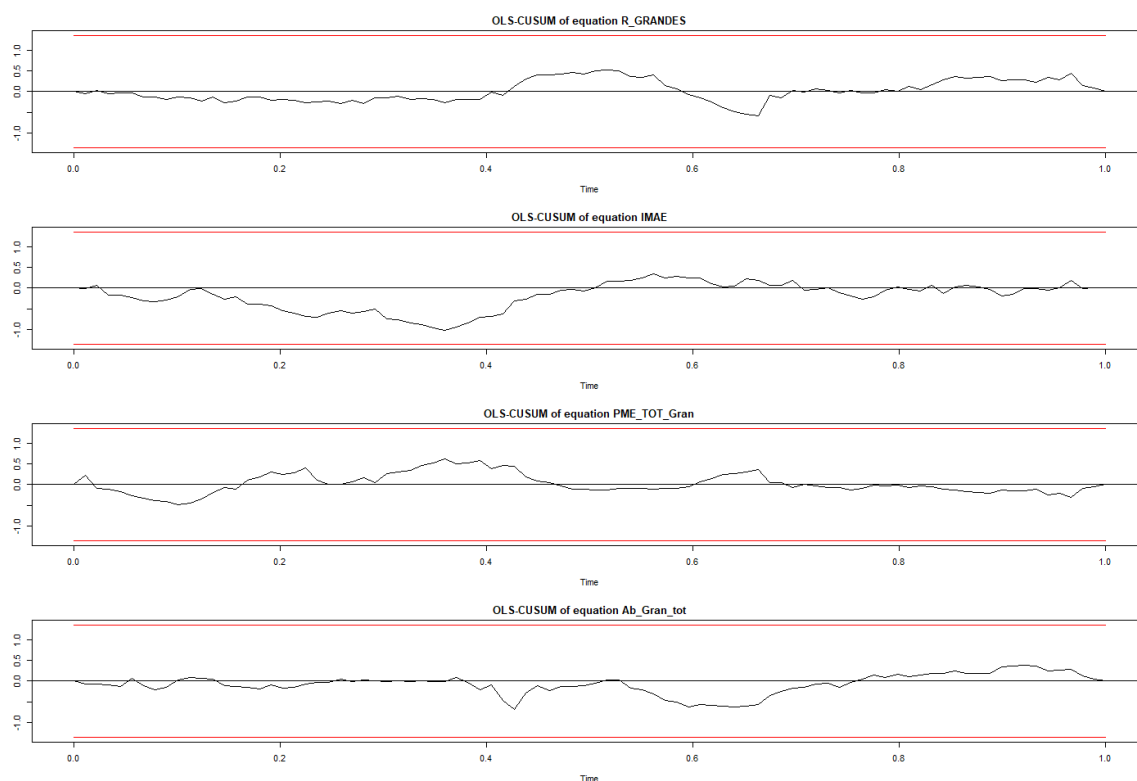
Para evaluar la presencia o ausencia de homocedasticidad del modelo se aplicó la prueba ARCH, obteniendo un p valor de 0.4737, con lo que no se rechaza la hipótesis nula de homocedasticidad.

Para analizar la estabilidad del modelo estimado se realiza la prueba de CUSUM. Como se observa en la figura 15 las variables se muestran estables en el tiempo y dentro de los intervalos de confianza estimados. El primer gráfico muestra el comportamiento de la demanda de los grandes consumidores, en el segundo gráfico se observa el comportamiento del IMAE, el tercer gráfico muestra el precio medio para la electricidad para los grandes consumidores y finalmente se muestra el comportamiento para el número de abonados en

este sector. Como se aprecia en todos los casos el comportamiento de las variables es aleatorio, algunas veces se encuentra por encima y otras por debajo de cero sin un patrón definido.

El no cumplimiento del supuesto de estabilidad de los coeficientes implica consecuencias serias por cuanto, en primer lugar, la estimación de los coeficientes produce resultados incorrectos, y en segundo lugar, porque las proyecciones resultan erróneas (BCCR,1996)

Figura 15. Prueba de estabilidad del modelo (OLS-CUSUM).



Fuente: Elaboración propia

Para evaluar la causalidad del modelo se realiza el test o prueba de causalidad de Granger, considerando la hipótesis nula de no causalidad (H_0 : R_GRANDES no tiene relación causal de Granger con respecto a IMAE, PME_TOT_Gran y Ab_Gran_tot), dado un valor $p = 0.000213$, se rechaza la hipótesis nula y por tanto se puede indicar que hay relación causal entre las variables explicativas y el consumo eléctrico de los grandes consumidores.

5.2.2 Demanda total grandes consumidores ICE.

El cuadro 10 se muestran los resultados estimados para el modelo VAR que estima el consumo eléctrico para los grandes consumidores del ICE. En la estimación se incluyeron las variables: consumo eléctrico de los grandes consumidores (I_GRANDES), índice mensual de actividad económica (IMAE), precio medio de la tarifa media tensión a para el caso del ICE (PME_GranICE) y cantidad de abonados del sector (Ab_TotICE). Para el caso de la estimación de la demanda eléctrica resultaron estadísticamente significativos los coeficientes asociados al rezago del consumo eléctrico y a la cantidad de abonados. El modelo cumple con la condición de estacionariedad dado que las raíces del polinomio se encuentran dentro del círculo unitario.

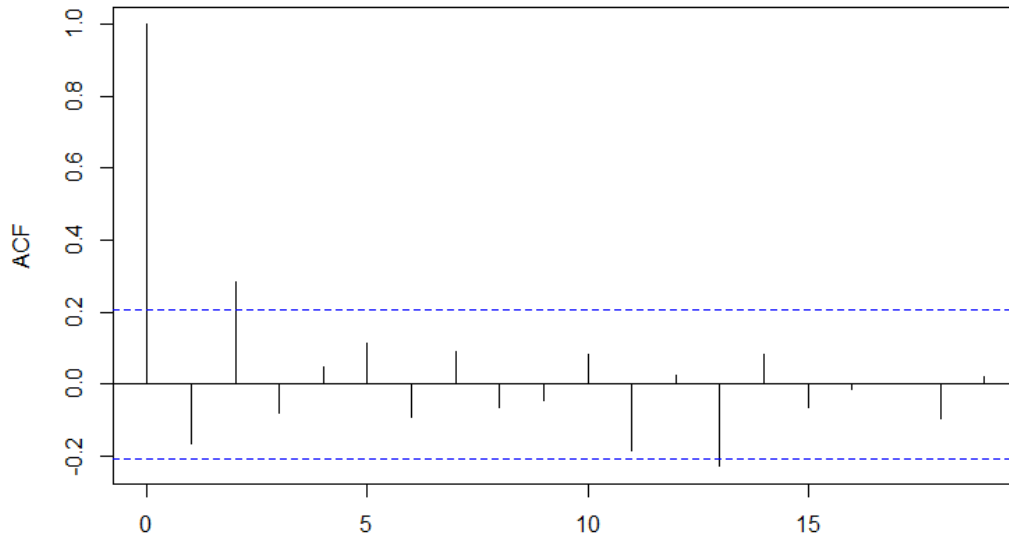
Cuadro 10. Resultados del modelo VAR para la demanda grandes consumidores ICE.

	Estimación	Error estándar	Valor-t	Pr(> t)
I_GRANDES.l1	0.8274	0.0630	13.1450	2,00E-16 ***
Ab_TotICE.l1	22.2241	8.2051	2.7090	0.0081 **

Fuente: Elaboración propia.

En el gráfico 9 se muestra el correlograma de los residuos para el modelo estimado, visualmente se observa que los residuos caen a cero rápidamente. Se realiza la prueba formal de Portmanteau multivariada para identificar la existencia o no de autocorrelación serial de orden superior, el valor p estimado es de 0.05244 como que existe evidencia estadística a favor de la hipótesis nula de no autocorrelación serial

Gráfico 9. Correlograma para los residuos del modelo.



Fuente: Elaboración propia

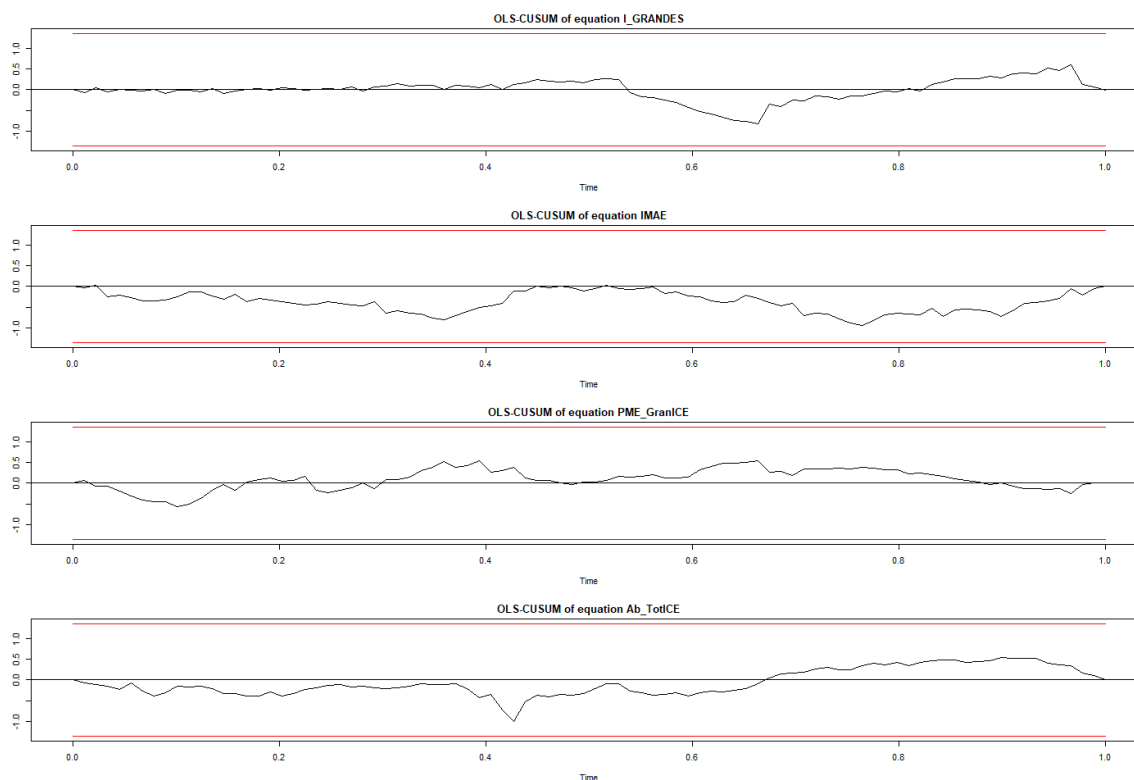
La normalidad de las variables y de los residuos del modelo se evalúa utilizando el estadístico Jarque Bera. Individualmente solo la variable del IMAE no rechaza la hipótesis nula de normalidad ($p= 0.2422$), a nivel global se rechaza la hipótesis nula de normalidad de los residuos del modelo ($p= 2.2e-16$). Por otra parte, la prueba de ARCH sobre heteroscedasticidad, mostro suficiente evidencia estadística para no rechazar el supuesto de homocedasticidad de los residuos ($p\text{-value}= 0.6458$)

La figura 16 muestra los resultados del análisis de estabilidad de modelo mediante la prueba de CUSUM, como se observa las variables se muestran estables en el tiempo y se ubican dentro de los límites de confianza calculados. El primer gráfico muestra el comportamiento de la demanda eléctrica para el sector en análisis, el segundo gráfico se observa el IMAE, el tercer gráfico muestra el precio medio de la electricidad y el cuarto el número de abonados. Tanto el IMAE como el precio medio no muestran un patrón aleatorio, en el primer caso las observaciones se encuentran por encima de 0 y en el segundo por debajo de dicho valor, una razón adicional para no resultar significativas estadísticamente en el modelo.

Adicionalmente se realizó la prueba de causalidad de Granger ($p=0.1905$) pero no se encontró suficiente evidencias estadística para rechazar la hipótesis nula de no causalidad entre la

variable consumo de electricidad para los grandes clientes del ICE y las variables explicativas.

Figura 16. Prueba de estabilidad del modelo (OLS-CUSUM).



Fuente: Elaboración propia

5.2.3 Demanda total sector residencial.

Los resultados del modelo VAR para la demanda del sector residencial se muestran en el cuadro 11. Las variables que se consideraron en la estimación fueron el total de abonados y el precio medio de la electricidad para el sector residencial, utilizando 3 rezagos. El modelo se estimó con las variables cuyos respectivos coeficientes resultaron estadísticamente distintos de cero. Se verifica la condición de estacionariedad dado que las raíces del polinomio son menores a uno.

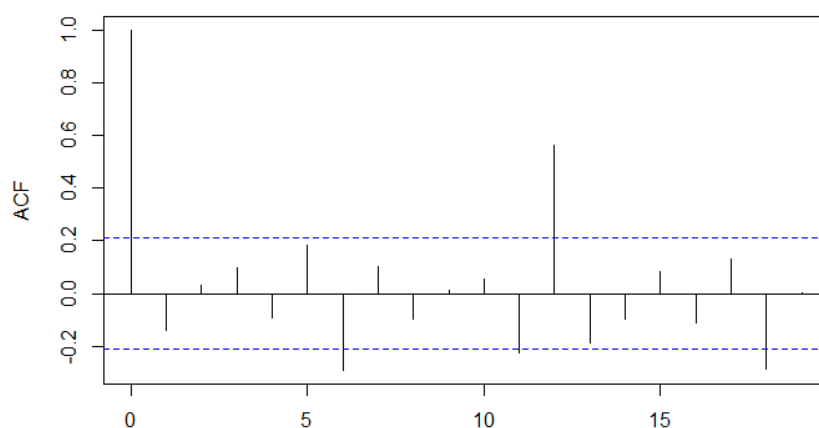
Cuadro 11. Resultados del modelo VAR para la demanda eléctrica sector residencial.

	Estimación	Error estándar	Valor-t	Pr(> t)
Ab_Res_Tot.l1	-2.881	1.062	-2.713	0.00811 **
Ab_Res_Tot.l2	3.082	1.062	2.903	0.00474 **
PMRes_tot.l2	-521.554	106.281	-4.907	4.54e-06 ***
Constante	70158.507	15150.262	4.631	1.33e-05 ***

Fuente: Elaboración propia.

El análisis de autocorrelación para los residuos del modelo se realiza mediante el correlograma del gráfico 10, adicionalmente se realiza la prueba formal de Portmanteau multivariada. Tanto el correlograma como la prueba formal muestran evidencia estadística a favor de la autocorrelación de orden superior. El valor $p=0.0366$ de la prueba de Portmanteau presenta evidencia estadística para rechazar la hipótesis nula de no autocorrelación serial.

Gráfico 10. Correlograma para los residuos del modelo.

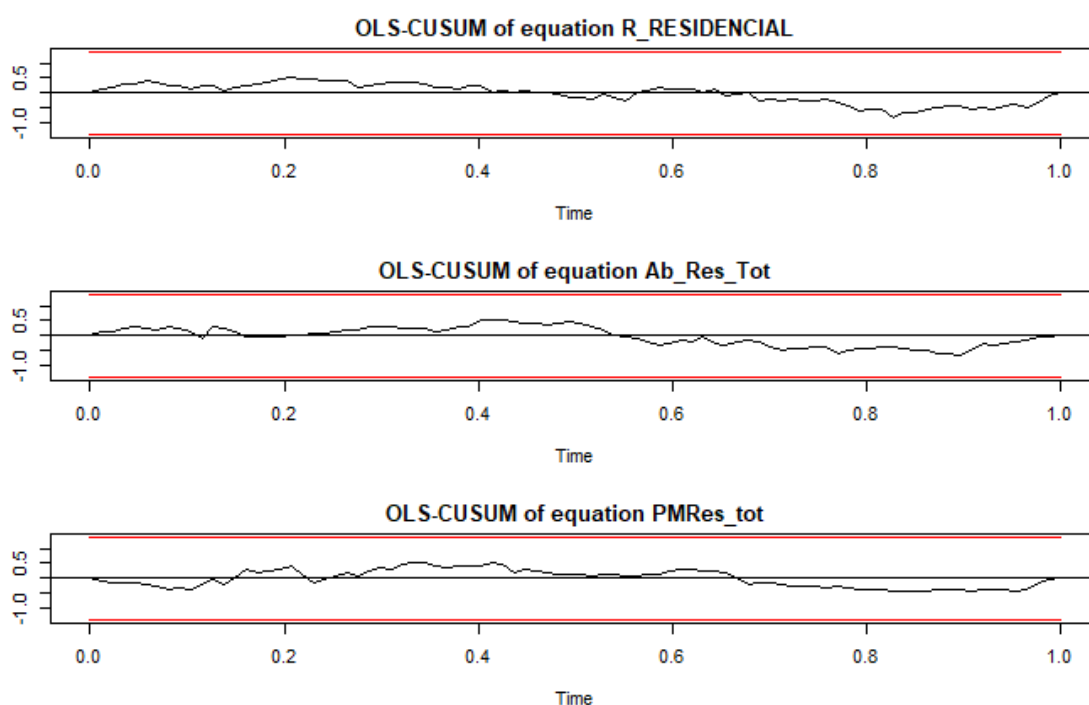


Fuente: Elaboración propia.

Mediante el estadístico Jarque Bera, se evaluó la normalidad de las variables a nivel individual y de los residuos del modelo, los valores p asociados al JB resultaron mayores al nivel de significancia del 5% con lo que se muestra evidencia a favor de la normalidad de las variables y de los residuales del modelo. El supuesto de homocedasticidad de los residuos se evalúa mediante la prueba de ARCH multivariada, obteniendo un p valor de 0.254. Los resultados muestran que se rechaza la hipótesis nula de heteroscedasticidad de los residuos.

La prueba gráfica para la estabilidad del modelo se muestra en la figura 17, que presenta los resultados para la prueba de CUSUM, las variables se muestran estables en el tiempo y se ubican dentro de los límites de confianza estimados. El primer gráfico muestra la demanda eléctrica, el segundo gráfico presenta el comportamiento de la cantidad de abonados y el tercer gráfico muestra el comportamiento del precio medio de la electricidad, en todos los casos se observa una senda aleatoria. La prueba de causalidad de Granger ($p=0.0033$) mostro evidencia estadística para rechazar la hipótesis nula de no causalidad entre la variable consumo de electricidad residencial y las variables de número de abonados y precio medio de la electricidad.

Figura 17. Prueba de estabilidad del modelo (OLS-CUSUM)



Fuente: Elaboración propia.

5.2.4 Demanda sector residencial del ICE

En el Cuadro 12 se presentan los resultados del modelo VAR para la demanda eléctrica del sector residencial del ICE. El modelo se estimó considerando la cantidad de abonados, el precio medio de la electricidad y la demanda eléctrica como variables, utilizando un rezago

de uno y dos períodos. Mediante un proceso de stepwise se eliminaron las variables cuyos coeficientes no resultaron estadísticamente significativos utilizando una matriz de restricción para el modelo estimado. Se verifica la condición de estacionariedad del modelo dado que las raíces características del polinomio son menores a cero. El coeficiente de bondad de ajuste R^2 es de 0.95 para el modelo seleccionado.

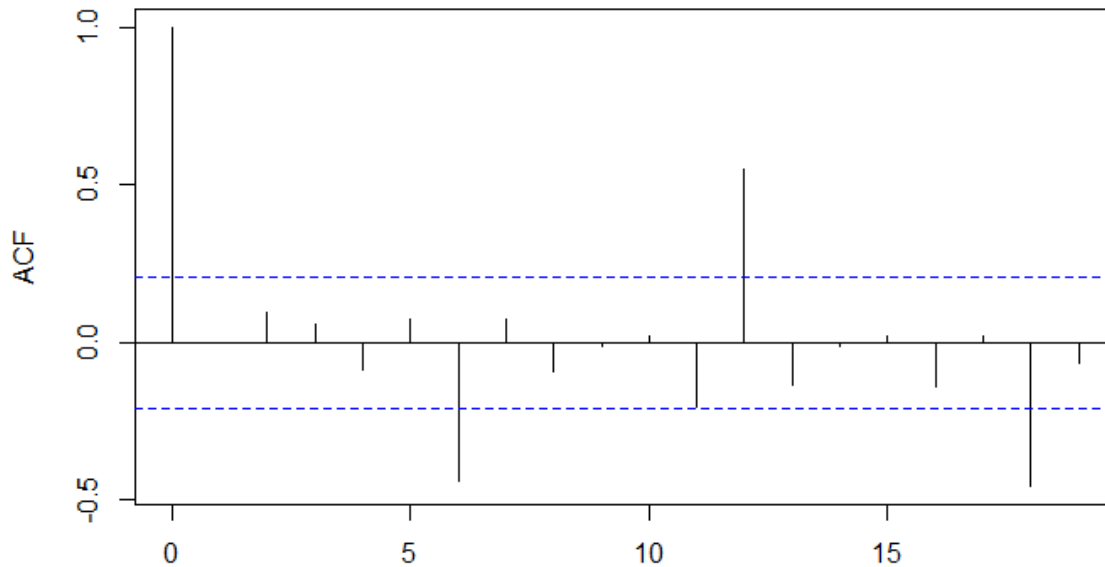
Cuadro 12. Modelo VAR demanda eléctrica sector residencial ICE.

	Estimación	Error estándar	Valor-t	Pr(> t)
I_RESIDENCIAL.I1	-0.0352	0.1097	-3.211	0.00189 **
Ab_Res_ICE.I1	-0.0059	0.0027	-2.195	0.03096 *
PMed_Res_ICE.I1	0.0001	0.5309	2.312	0.02326 *
Ab_Res_ICE.I2	0.0069	0.0027	2.545	0.01281 *
PMed_Res_ICE.I2	-0.0001	0.5423	-3.715	0.00037 ***
Constante	-463.50	51.67	-8.972	8.18e-14 ***

Fuente: Elaboración propia.

El análisis de autocorrelación para los residuos del modelo se realiza mediante el correlograma del gráfico 11 y la prueba formal de Portmanteau multivariada. Tanto el correlograma como la prueba formal muestran evidencia estadística a favor de la autocorrelación de orden superior. El valor $p = 0.004948$ de la prueba de Portmanteau presenta evidencia estadística para rechazar la hipótesis nula de no autocorrelación serial; el correlograma se observan algunos residuos que salen del intervalo.

Gráfico 11. Correlograma para los residuos del modelo.

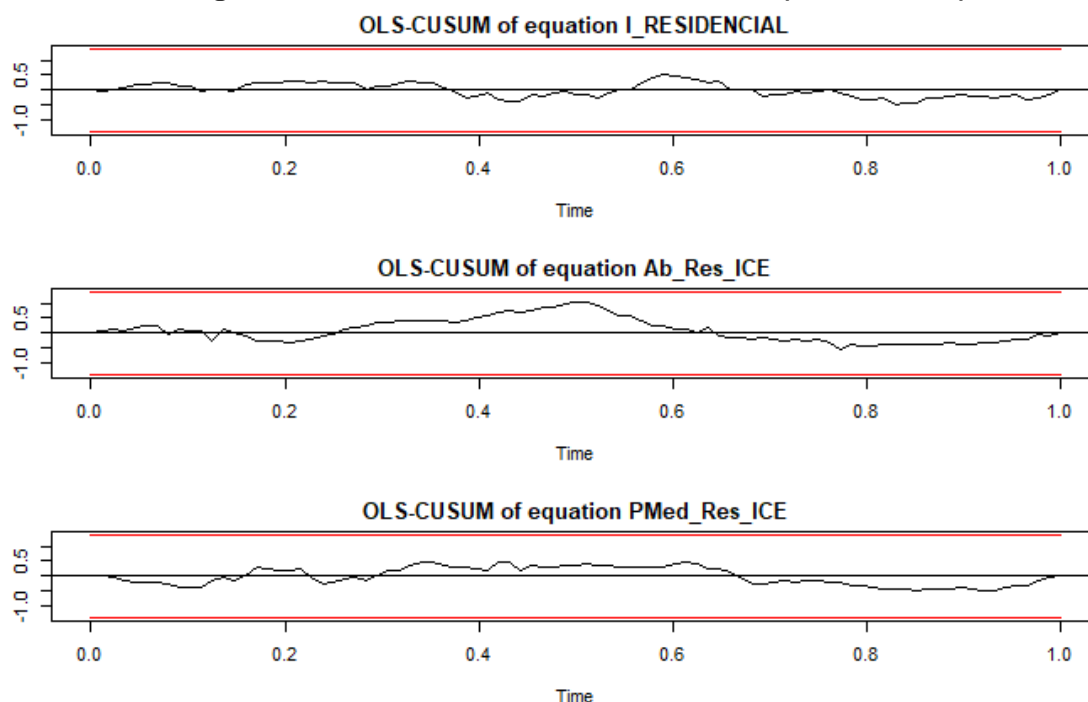


Fuente: Elaboración propia.

El análisis de la normalidad se realiza mediante el estadístico JB, a nivel individual solo la variable número de abonados rechaza la hipótesis nula de normalidad ($p=0.0001887$), a nivel del modelo el valor p estimado para el estadístico JB es de 0.0001887 con lo que se rechaza la hipótesis nula de normalidad de los residuales. El test de ARCH multivariado muestra presencia de heteroscedasticidad de los residuos para rezagos menores de 8.

La estabilidad de modelo se verifica mediante la prueba de CUSUM, que se muestra en la figura 18, se corrobora que las variables se muestran estables en el tiempo y se ubican dentro de los límites de confianza calculados, la primera gráfica se refiere a la demanda eléctrica, la segunda corresponde al número de abonados del sector residencial y la tercera al comportamiento del precio promedio de la electricidad para el sector residencial. La prueba de causalidad de Granger muestra evidencia estadística suficiente para rechazar la hipótesis nula de no causalidad entre las variables explicativas y la variable dependiente del modelo ($p=2.029e-05$).

Figura 18. Prueba de estabilidad del modelo (OLS-CUSUM).



Fuente: Elaboración propia.

5.3 Comparación de modelos

Para realizar una comparación de los modelos se utiliza el error cuadrático medio, cuyo propósito es medir la diferencia o error entre dos conjuntos de datos. Para este caso se comparan los resultados de los modelos ARIMA y VAR considerando una proyección a 6 meses y los valores reales observados para ese período. La proyección se realizó para el segundo semestre del 2019.

El cuadro 13 muestra un resumen del Error Cuadrático Medio (RSME, por sus siglas al inglés) obtenido para las proyecciones realizadas tanto con el modelo ARIMA como para el VAR. Al analizar las proyecciones de demanda para los grandes consumidores a nivel país y para los clientes del ICE, se observa como este es significativamente menor en las proyecciones realizadas mediante los modelos ARIMA. De igual manera, cuando se observa el caso de la demanda estimada para el sector residencial el RSME es menor cuando se utiliza un modelo ARIMA. Cabe indicar que para el caso del sector residencial la diferencia entre

el error cuadrático medio obtenido en ambos modelos es mucho menor que la observada para el grupo de grandes consumidores.

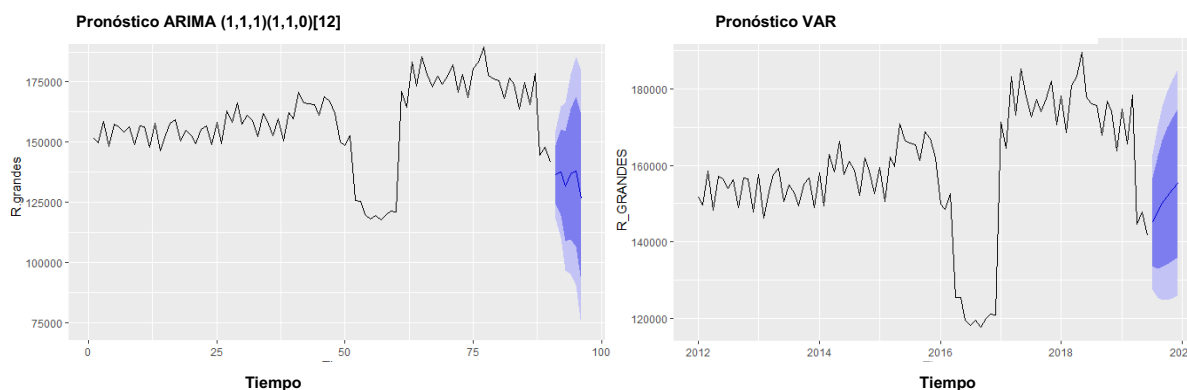
Cuadro 13. Error cuadrático medio para la proyección de la demanda eléctrica.

	<i>Grandes Consumidores</i>		<i>Sector Residencial</i>	
	Total País	ICE	Total País	ICE
ARIMA	8 708,05	1 382,82	6 006,05	2 705,47
VAR	22 121,06	8 513,86	6 461,69	3 510,43

Fuente: Elaboración propia.

En la figura 19 se compara la proyección realizada con el modelo ARIMA (izquierda) y VAR (derecha) para la demanda eléctrica del país, el caso de los grandes consumidores. Como se observa el modelo ARIMA captura mejor la tendencia y muestra variabilidad en la estimación, por su parte, el modelo VAR sobre estima los valores observados reales y genera una proyección creciente de la demanda.

Figura 19. Proyección de la demanda eléctrica grandes consumidores total país.

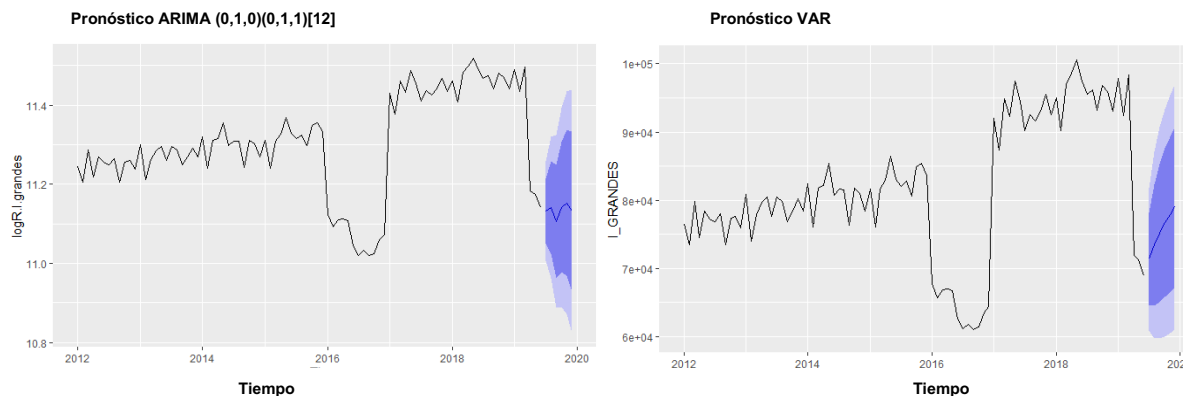


Fuente: Elaboración propia.

Las proyecciones para el caso de la demanda eléctrica de los grandes consumidores del CIE se aprecian en la figura 20. Al igual que en la proyección para la demanda nacional, el modelo VAR reproduce una proyección con tendencia creciente sobre estimando los valores observados. El modelo ARIMA muestra cierta variabilidad en la estimación lo que es más

acorde con el comportamiento observado de la demanda eléctrica para este grupo de consumidores.

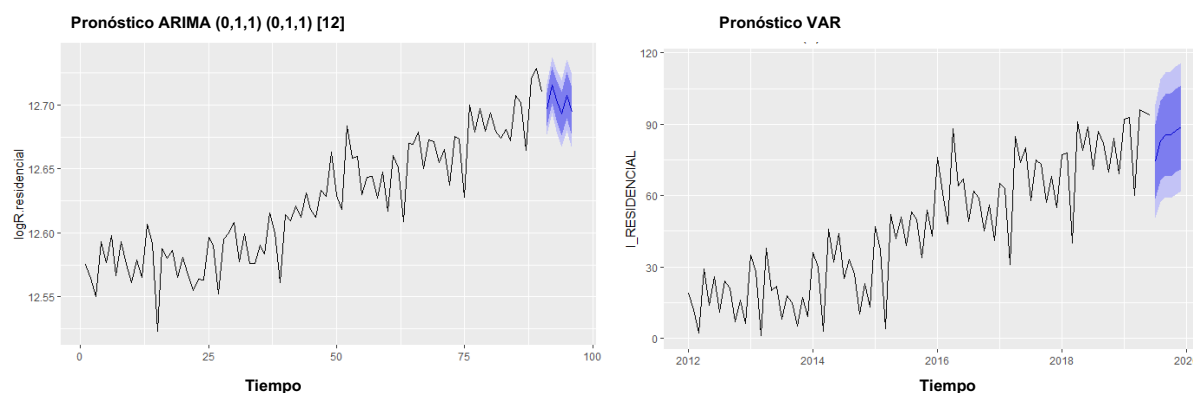
Figura 20. Proyección de la demanda eléctrica grades consumidores ICE.



Fuente: Elaboración propia.

La figura 21 muestra las proyecciones para la demanda eléctrica del sector residencial del país, la proyección a 6 meses estimó un menor RSME para el modelo ARIMA, que sigue mostrando un mejor ajuste comparado al VAR. El modelo VAR nuevamente tiende a estimar un crecimiento en la demanda.

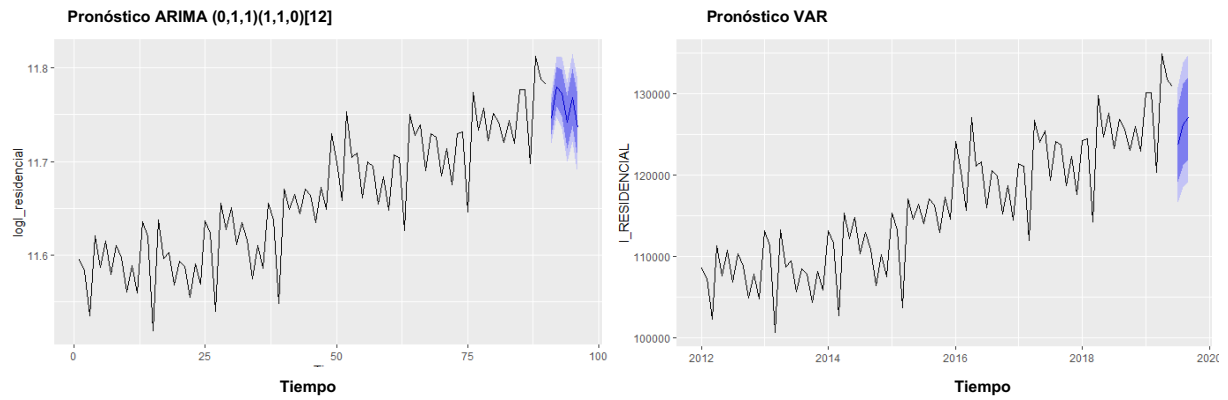
Figura 21. Proyección de la demanda eléctrica sector residencial total país.



Fuente: Elaboración propia.

La proyección para el sector residencial del ICE se muestra en la figura 22, de igual manera el modelo VAR tiende a estimar una proyección creciente de la demanda eléctrica.

Figura 22. Proyección de la demanda eléctrica residencial del ICE.



Fuente: Elaboración propia.

5. CONCLUSIONES

La estimación y proyección de la demanda eléctrica con el menor margen de error posible es deseable para la toma de decisiones en diferentes aspectos del quehacer regulatorio, con implicaciones en los ámbitos económicos, sociales y ambientales. Una estimación inadecuada de la demanda, por ejemplo, para fines tarifarios, puede hacer una transferencia excesiva de renta de unos agentes económicos a otros, generando una pérdida de bienestar y eficiencia considerable. La demanda eléctrica es un componente importante para el cálculo de la tarifa en diferentes sectores, una buena estimación contribuye a aproximar una tarifa cercana a los costos por kWh de brindar el servicio evitando subestimaciones o sobreestimaciones excesivas y dando señales correctas a los agentes económicos.

Este trabajo utiliza dos tipos de técnicas para estimar y proyectar la demanda eléctrica con el propósito de identificar la que disminuya las diferencias entre los valores observados y proyectados, es decir, encontrar el mejor modelo para efectos de pronóstico. Los modelos ARIMA se utilizaron para explicar y proyectar el comportamiento de la demanda eléctrica mediante sus valores pasados; por otra parte, los modelos VAR buscan lograr el mismo objetivo, pero mediante un sistema de ecuaciones considerando diferentes variables que pueden incidir en la demanda eléctrica y los rezagos de esas variables. Ambas técnicas fueron aplicadas a la demanda de los clientes residenciales y a la demanda de los clientes cuya tarifa es denominada “media tensión a” referida a los grandes consumidores. El ejercicio se realiza para el total de clientes a nivel país y para los clientes del ICE.

La selección de la forma funcional del modelo ARIMA se realizó considerando 3 métodos distintos. El primer proceso consistió en seguir las etapas propuestas por Pankratz (1983) siguiendo el modelo de Box Jenkins para el análisis de series de tiempo. Lo anterior implica la identificación de un modelo ARIMA basado en el análisis y estudio completo de la serie de tiempo, para entender su estructura y comportamiento durante el periodo analizado, identificando sus condiciones de estacionariedad, estacionalidad y estructura (ARIMA o SARIMA) basado en el análisis de los autocorrelogramas; posteriormente se estima y verifica el modelo seleccionado. Para el segundo método se utilizó la función `auto.arima` del software estadístico R, que realiza una serie de combinaciones de modelos ARIMA y selecciona aquel

con menor indicador de ajuste medido por el AIC. El último método consistió en utilizar el paquete Tramo-Seat del software econométrico Eviews para la identificación de un modelo.

Una conclusión importante a tener en cuenta con el ejercicio anterior, se refiere a la precaución y una llamada de atención al utilizar la función `auto.arima`. Para todos los modelos identificados mediante esta función, es claro que prioriza aquel modelo con un menor valor de ajuste (AIC), sin tener en cuenta aspectos estructurales de la serie. Por ejemplo, `auto.arima` puede obviar en muchos casos el hecho que la serie es no estacionaria y seleccionar un proceso ARIMA que no incluye diferenciación alguna de la serie. La mayor parte de las series económicas son no estacionarias, de allí la importancia de diferenciarlas para que su media y varianza no cambien en el tiempo. Otro aspecto a tener en cuenta al menos para este caso, es que la función `auto.arima` tiende a seleccionar un mayor número de autoregresivos y medias móviles que los que se identifican mediante las funciones autocorrelación y autocorrelación parcial. Por estas razones, si no se revisan rigurosamente los resultados de este paquete se puede conducir a estimaciones y formas funcionales erróneas.

Con respecto a la función Tramo-Seat del programa econométrico Eviews se puede decir que es una buena guía para la revisión o identificación a priori de un modelo, pues los modelos seleccionados mediante este programa fueron cercanos o similares a los seleccionados siguiendo un proceso “manual”, sin embargo, no necesariamente correspondían a un modelo con el menor AIC. Un aspecto que es un valor agregado de Tramo-Seat, a pesar de que no se hizo necesario considerar en estos modelos, es que permite identificar días festivos² o aspectos que pueden requerir un análisis de intervención.

De manera general, cualquier análisis riguroso de series de tiempo con propósitos de pronóstico debe seguir un estudio y análisis detallado de la serie de tiempo, identificar si cumple los supuestos de estacionariedad y realizar las transformaciones que se requieran antes de iniciar la etapa de identificación, evaluar cambios estructurales u observaciones

² Tramo/Seat permite identificar si días festivos como por ejemplo Semana Santa, Navidad u otros tienen un efecto no observado sobre la serie. De existir este tipo de efectos se requiere del análisis de intervención en donde particularmente se ajusta nuevamente el modelo mediante el uso de variables dummy para considerar el efecto de los días festivos.

atípicas de la serie. Una vez realizado este primer análisis y realizadas las transformaciones necesarias, en la etapa de identificación corresponde evaluar diferentes procesos generadores AR, MA, ARMA o SARMA, partiendo del análisis de las funciones de autocorrelación y autocorrelación parcial, para posteriormente evaluar la significancia estadística de los coeficientes y el cumplimiento de los supuestos de los residuales del modelo.

En cuanto a los modelos VAR, es importante reconocer su poder y utilidad práctica pues permiten incorporar en la estimación otras variables que pueden incidir en la variable a predecir, que un modelo ARIMA no observa. Sin embargo, a pesar de que se consideraron en cada caso un sistema de ecuaciones incluyendo aquellas variables que explican el consumo o la demanda eléctrica desde el punto de vista teórico, algunas de ellas y sus rezagos no resultaron estadísticamente significativas.

Para este ejercicio los modelos VAR generaron mayores RSME con respecto al modelo ARIMA. Sin embargo, hay algunas consideraciones a manera de reflexión que se pueden tener en cuenta en esta estimación y que pueden haber incidido en para este caso particular el modelo VAR no genera mejores resultados:

- En primer lugar, se utilizaron los precios medios de la electricidad. Para el caso costarricense los precios no necesariamente reflejan variaciones mensuales o en periodos menores, esto debido a que cada fijación tarifaria se realiza para periodos de un año y se ajusta cada 3 meses por efecto de un componente denominado Costo Variable de Combustible, la falta de variabilidad en el precio puede conducir a que el VAR no logre capturar realmente el efecto de esta variable con respecto a las demás, adicionalmente el efecto de sus rezagos en el corto plazo pueden no ser observables.
- El número de abonados si bien varía porque se incorporan nuevos usuarios al sistema, no depende de los cambios en el precio, pues el usuario no tiene la capacidad de elegir un distribuidor particular para comprar electricidad.
- El precio del GLP no mostró coeficientes estadísticamente significativos, a pesar de ser un bien que se considera sustituto y que puede explicar otras variables como el IPC o el IMAE.

- El IMAE y el IPC, a pesar de que desde el punto de vista teórico pueden relacionarse con la demanda eléctrica, sus coeficientes no mostraron efecto importante en las estimaciones, con excepción del IMAE que se incorporó en la demanda eléctrica para el caso de los grandes consumidores.

Los modelos ARIMA se consideran más robustos para estimar y pronosticar el comportamiento de la demanda eléctrica, a pesar de mostrar en la mayoría de los casos problemas de autocorrelación de los residuos. A pesar de lo anterior, las pruebas de estabilidad y causalidad fueron aplicadas con resultados satisfactorios, lo que permite concluir que pueden ser utilizados para fines de pronóstico. Se recomienda continuar explorando otro tipo de técnicas estadísticas y econométricas que permitan comparar los resultados de los modelos expuestos en este trabajo con otros, con el propósito de buscar estimaciones cada vez más precisas.

6. REFERENCIAS

Adrian Trapletti and Kurt Hornik (2013). tseries: Time Series Analysis and Computational Finance. R package version 0.10-32.

Alan Agresti (2013): Categorical data analysis," Statistical Papers, Springer, vol. 57(3), pages 849-850, September.

Alan Pankratz, "Forecasting with Univariate Box-Jenkins Models", New York: John Wiley and Sons, 1983.

Camacho, O. S. A. (2006). Aspectos computacionales de algunos métodos de ajuste paramétrico de modelos. 164.

Club Español de la Energía. (2014). El Sector Energético Español y su Aportación a la Sociedad.

Barrientos Marín, Jorge y Toro Martínez, Mónica (2017). Análisis de los fundamentales del precio de la energía eléctrica: evidencia empírica para Colombia. Revista de economía del Caribe N°. 19. ISSN: 2145-9363 (on line).

E. Bompard, G. Ciwei, R. Napoli and F. Torelli. (2007) Dynamic price forecast in a competitive electricity market. The Institution of Engineering and Technology 2007.

Espasa, Antoni y Pérez Espartero, Ana. (2008). Forecasting with dynamic regression models. 25 noviembre, 2008.

Espínola Vilchez, Rosario. (2004) Predicción a corto plazo de los precios de la energía eléctrica. Universidad de Castilla – La Mancha. Escuela Técnica Superior de Ingenieros Industriales Departamento de Ingeniería Eléctrica, Electrónica y Automática. Octubre 2004.

García-Ascanio, Carolina y Mate, Carlos. (2010). Electric power demand forecasting using interval time series: A comparison between VAR and iMLP. Volume 38, Issue 2, February 2010, Pages 715-725. Recuperado de: <https://doi.org/10.1016/j.enpol.2009.10.007>

Gujarati, D. N., (2010). Econometría: Damodar N. Gujarati. (5a.ed.--.). México: McGraw Hill.

Hannan, E.J. y B.G. Quinn (1979). "The Determination of the Order of an autoregression", Journal of the Royal Statistical Society B 41, pp. 190-195.

Instituto Costarricense de Electricidad, 2017. Planificación y desarrollo eléctrico proceso expansión del sistema 2016-2035. Proceso Expansión del Sistema de la Dirección de Planificación y Desarrollo Eléctrico, Gerencia Electricidad, Instituto Costarricense de Electricidad. Mayo 2017 San José, Costa Rica.

Leiva Sandoval, Carlos. (2009). Las elasticidades de la energía comercial en Costa Rica. Ministerio del Ambiente, Energía y Telecomunicaciones. Dirección Sectorial de Energía. San José, Costa Rica.

Orellana Romero. (2012). Modelación y Pronóstico de la Demanda de Energía Eléctrica de Mediano Plazo de El Salvador. Universidad de El Salvador Facultad de Ciencias Naturales y Matemáticas. Escuela de Matemática. Maestría en Estadística. Tesis para optar por el grado de Maestría en Estadística.

Pastor Martínez, Alfredo y Juan Ruiz, Jesús (2015). Modelo de predicción de la demanda eléctrica mediante regresión dinámica. Universidad Politécnica de Madrid. Escuela Técnica Superior de Ingenieros Industriales. Julio 2015, Madrid. Recuperado de http://oa.upm.es/43848/1/TFG_ALFREDO_PASTOR_MARTINEZ.pdf

Pfaff, B. (2008) Analysis of Integrated and Cointegrated Time Series with R. Second Edition. Springer, New York. ISBN 0-387-27960-1

Rahimi Gollou, Abbas & Ghadimi, Noradin. (2017). A new feature selection and hybrid forecast engine for day-ahead price forecasting of electricity markets. *Journal of Intelligent & Fuzzy Systems*. 32 (2017) 4031–4045.

Vanderkam, D., Allaire, J., Owen, J., Gromer, D., and Thieurnel, B.: dygraphs: Interface to “Dygraphs” Interactive Time Series Charting Library, available at: <https://CRAN.R-project.org/package=dygraphs>, R package version 1.1.1.6, 2018.

Vilar, Juan., Aneiros, Germán., y Paula Raña. (2017). Prediction intervals for electricity demand and price using functional data. *Electrical Power and Energy Systems* 96 (2018) 457–472 <http://dx.doi.org/10.1016/j.ijepes.2017.10.010>.

Weron Rafal. (2014). Electricity Price forecasting: A review of the state of the art with a look into the future *International Journal of Forecasting* 30(2014)1030–1081. *International Journal of Forecasting*. <http://dx.doi.org/10.1016/j.ijforecast.2014.08.008>

Wickham, H. (2009) *Ggplot2: Elegant Graphics for Data Analysis*. 2nd Edition, Springer, New York. <https://doi.org/10.1007/978-0-387-98141-3>